

;; PRIOR FILING DATE: 2004-05-04
;; NUMBER OF SEQ ID NOS: 107
;; SOFTWARE: PatentIn version 3.3
;; SEQ ID NO 61
;; LENGTH: 169495
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: modified_base
;; LOCATION: (70072)..(70171)
;; OTHER INFORMATION: a, c, g, t, unknown or other
;; NAME/KEY: modified_base
;; LOCATION: (139457)..(157244)
;; OTHER INFORMATION: a, c, g, t, unknown or other
US-11-121-086-61

Query Match 1.2%; Score 42.4; DB 7; Length 169495;
Best Local Similarity 48.7%; Pred. No. 5.5;
Matches 115; Conservative 0; Mismatches 121; Indels 0; Gaps 0;

Qy 2389 GGTGTTAGTAGATCCCTGTTGGTAAATGAGGCGGTGATGGTGAAGCTGGTCTTT 2448
Db 67566 GTGGTCTCTGATGCTGTTGATATCTGTTGGTGGTCTTTGCTTTGCTGCT 67507
Qy 2449 ATTGCAAGCTGTAAAGCCCTTGAAACTTTGAACTAATGATTAATGCAATTCGAGCT 2508
Db 67506 ACTGTTGGTGTACTCTGTTGTTGACGCTGTTGATGCTGCTGTTGATGAT 67447
Qy 2509 GGGCTCTAGAGAGAGAGGTTGATTAATACTTCTGTTGAAAGTCCGTTGCTGATATGTTA 2568
Db 67446 GCTGTGATGATGCTGTTGTTGATGCTGTTGTTGATGATGCTGTTGTTGATGAT 67387
Qy 2569 AAGAGGCAAGGTGTAAGATCTAGCTGCTGGTGGCAATTTGACTGATTTCACT 2624
Db 67386 ATTGCTGCTTTTGGTGTGTTGCTGCTGCTGCTGATGAGTGGTGGTGTCTGCTCT 67331

RESULT 3
US-10-750-185-40551/c
; Sequence 40551, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 40551
; LENGTH: 1199
; TYPE: DNA
; ORGANISM: Bovine
US-10-750-185-40551

Query Match 1.1%; Score 42; DB 6; Length 1199;
Best Local Similarity 50.5%; Pred. No. 0.26; Indels 0; Gaps 0;
Matches 102; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

Qy 2679 GATTCTTCTATGATCTGATGCTACACATAGATGATTAATGATGCAATTTTCA 2738
Db 2732 GTTCTCTTTTGGAACTAAGCTTCAAGTGTAAATGCTGAGAAATACACATTAATGT 213
Qy 2739 TATATCTGATTTCTCAAAATATGCTTTGTTGAGCTAAGAACATATGTTCCCATTTAA 2798

Db 212 TTTCCTTTCTTAATTTTCAATTTCCCTGTTTCTGCTCTCAATTCAGAGATGTTTCCGT 153
Qy 2799 TACATGCCCAAGTGTACCAAGATTAACAGTTGCTGAGTAATTTGACTAATATG 2858
Db 152 TTCAATGCCCTAATTTTATATCTGAGTTATCAATTTGGTAATCATATTTTAACTTTGCA 93
Qy 2859 CTGCTGAATTTTGTGATCAAA 2880
Db 92 GAGTTCTTTTATTTCTCTTAA 71

RESULT 4
US-11-121-086-75/c
; Sequence 75, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: FOULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 75
; LENGTH: 161874
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-75

Query Match 1.1%; Score 40.6; DB 7; Length 161874;
Best Local Similarity 47.8%; Pred. No. 17;
Matches 118; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

Qy 2753 TCAAAATATGCTTTGTTGAGCTAAGAACATAGTTCCCATTAATACATGTCCCAAA 2812
Db 49935 TCATTAATGATTTATTAATCTTAAGATTATTAATTAATTAATCTTTTCCCTTTAGCCCA 49876
Qy 2813 GTTGATCAAGATTAACAAATTTGCTGAGTAATTTTCACTAATATGCTGTAATTTT 2872
Db 49875 AATGTTTCCATGTCACCTGACAGACAGAAATTAATTAATTAATTAATTAATTAAT 49816
Qy 2873 TGATCAATCTGTAGACAGAAATGTAATTTCACTCTCAATTTCTGTTAGATAAGT 2932
Db 49815 TGATGAATAATGATGATTTAATGTTAATGAGACATGCAAACTTTTCTGTATATGATCCA 49756
Qy 2933 AGGATTAGAGATTGCCCTTATGCTGCTTGTGCCAATTTTCTTCCCTGATTTTCTT 2992
Db 49755 TATAGTAACATTTTCACTTATGAGATCTACAACTTATTTGTGCACTACTATATCTC 49696
Qy 2993 TTGCATT 2999
Db 49695 TGCCATT 49689

RESULT 5
US-10-829-826B-21
; Sequence 21, Application US/10829826B
; Publication No. US20050266397A1
; GENERAL INFORMATION:
; APPLICANT: Ecker, David J.
; APPLICANT: Hostadler, Steven A.
; APPLICANT: Sampath, Rangarajan
; APPLICANT: Blyn, Lawrence B.
; APPLICANT: Hall, Thomas A.
; APPLICANT: Massire, Christian
; TITLE OF INVENTION: METHOD FOR IDENTIFICATION OF CORONAVIRUSES
; FILE REFERENCE: IBIS0075-100 (IBIS-0058US)
; CURRENT APPLICATION NUMBER: US/10/829,826B
; CURRENT FILING DATE: 2004-04-22

```

; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 31028
; TYPE: DNA
; ORGANISM: Coronavirus
; US-10-829-826B-21

```

Query Match	1.1%;	Score 40.2;	DB 6;	Length 31028;
Best Local Similarity	51.4%;	Pred. No. 7.1;		
Matches	93;	Conservative	0;	Mismatches 88;
				Indels 0;
				Gaps 0;

QY	2383	GAGCAGAGGTTTAGTGTGATGATCCGTGTGTGAACAATGTATAGGCCGATGATGAGCGCTGT	2444
Db	3700	GATGCTATGTTCTTTTATGTGTGATGATGTTGTCTCAATGTGTGCAAGTGTGTGATGCTATNG	3755
QY	2443	GTCCTTTATTCAGAGAAGCTGTATAGACCCCTGTGAAAATTGTAAACTATATGATTTATGCAATT	2502
Db	3760	GTACTTATGTATGTTGTATGTATGTCACATTATACAGCCCACTTTGCTCTTAAAGATAAGTTGTTT	3815
QY	2503	CGAGCTGGGGTCTCAGAGAGTAGGCGTTGATGAATACTAGTTGAAATGTCCGTTGCTGAT	2566
Db	3820	TGTGCAATTTATTAAGCGTAGTGATGTATTAAGACAGCTGTGTGTGTGATGATTTATGAT	3875
QY	2563	A 2563	
Db	3880	A 3880	

RESULT 6
US-10-829-826B-22

```

1  APPLICANT: Ecker, David J.
2  APPLICANT: Hotelstädler, Steven A.
3  APPLICANT: Sampath, Rangarajan
4  APPLICANT: Blyn, Lawrence B.
5  APPLICANT: Hall, Thomas A.
6  APPLICANT: Maesire, Christian
7  TITLE OF INVENTION: METHOD FOR IDENTIFICATION OF CORONAVIRUSES
8  FILE REFERENCE: 1R150075-100 (DIBIS-0058US)
9  CURRENT APPLICATION NUMBER: US/10/829,826B
10 CURRENT FILING DATE: 2004-04-22
11 NUMBER OF SEQ ID NOS: 105
12 SOFTWARE: FastSeq for Windows Version 4.0
13 SEQ ID NO 22
14 LENGTH: 31028
15 TYPE: DNA
16 ORGANISM: Coronavirus
17 IS-10-829-826B-22

```

Query Match	1.1%	Score 40.2;	DB 6;	Length 31028;
Best Local Similarity	51.4%;	Pred. No. 7.1;		
Matches	93;	Conservative	0;	Mismatches 88;
			Indels	0;
			Gaps	0;

QY	2383	GAGCAGCTGTTTACTGTAAGATCCCTGTTGGTAAACAATGTAGGCCCGGATGCTGAGCTCGT	2442
Db	3700	GAGCTATGTTCTTTTATGATGATGTTGTCACATGTGTCCAACTGTGGAGCTATG	3755
QY	2443	GCTCTTATTCAGAGAAGCTGTAAAGACCCCTGTGAAAACCTTGAAACTATGATTTATGCATTT	2502
Db	3760	GTAATTATGATGTGTGATGTGCCATTTTACAGCCCACTTGTGCTTTAAAGATTAAGTGTGTT	3819
QY	2503	CGAGCTGGGGCTTCAGAGAGTAGCGTTGATGAAACTACTGTTGAATGTCCGTTCTGAT	2562
Db	3820	TGTCATTTTATTAAGACGTAGGTGTATTAAGACGCTGTGTGTGATGTTATATGAT	3875
QY	2563	A 2563	
Db	3880	A 3880	

RESULT 7
US-10-829-826B-26
; Sequence 26, Application US/10829826B
; Publication No. US20050266397A1

```

1  APPLICANT: Eckert, David J.
2  APPLICANT: Hofstadler, Steven A.
3  APPLICANT: Sampath, Rangarajan
4  APPLICANT: Blynn, Lawrence B.
5  APPLICANT: Hall, Thomas A.
6  APPLICANT: Maesite, Christian
7  TITLE OF INVENTION: METHOD FOR IDENTIFICATION OF CORONAVIRUSESE
8  FILE REFERENCE: IBIS0075-100 (IBIS-0058MS)
9  CURRENT APPLICATION NUMBER: US/10/829,826B
10 CURRENT FILING DATE: 2004-04-22
11 NUMBER OF SEQ ID NOS: 105
12 SOFTWARE: FastSeq for Windows Version 4.0
13 SEQ ID NO 26
14 LENGTH: 31028
15 TYPE: DNA
16 ORGANISM: Coronavirius
17 US-10-829-826B-26

```

Query Match	1.1%;	Score 40.2;	DB 6;	Length 31028;
Best Local Similarity	51.4%;	Pred. No. 7.1;		
Matches	93;	Conservative	0;	Mismatches 88;
			Indels	0;
			Gaps	0;

[illegible]

RESULT 8
US-10-82

```

Sequence 24, Application US/10829826B
Publication No. US20050266397A1
GENERAL INFORMATION:
APPLICANT: Ecker, David J.
APPLICANT: Horstadler, Steven A.
APPLICANT: Sampath, Rangarajan
APPLICANT: Blyn, Lawrence B.
APPLICANT: Hall, Thomas A.
APPLICANT: Maesite, Christian
TITLE OF INVENTION: METHOD FOR IDENTIFICATION OF CORONAVIRUSES
FILE REFERENCE: IB50075-100 (DIBIS-0058US)
CURRENT APPLICATION NUMBER: US/10/829,826B
CURRENT FILING DATE: 2004-04-22
NUMBER OF SEQ ID NOS: 105
SOFTWARE: PatcSeq for Windows Version 4.0
SEQ ID NO 24
LENGTH: 31100
TYPE: DNA
ORGANISM: Coronavirus
US-10-829-826B-24

```

Query Match	1.1%	Score 40.2;	DB 6;	Length 31100;
Best Local Similarity	51.4%;	Pred. No. 7.1;		
Matches 93;	Conservative 0;	Mismatches 88;	Indels 0;	Gaps 0;

2383 GAGACAGTGTTCAGTAGATCCTGTTGGTAACAATGTAGGCCGTGATGTCGTCGT 2442

Db 3700 GATGCTATGTTCTTTTATGAGTGTGTCACATGTCGCAAGTGTGAGTCTATG 3759
Qy 2443 GTCTTATTTGAGAGAGCTGTAGAACCTCTGTGAAAATTGAACTATGATGATCAAT 2502
Db 3760 GTACTATTGATGTTGATGTGCGCATTTACAGCCCACTTGTCTTTAAGATAAGTGT 3819
Qy 2503 CGAGCTGGGGTCTCAGAGATGACGCTTGATGAAACTACTGTTGAAAGTCCGTTGCTGAT 2562
Db 3820 TGTGATTTATTAATAAGCTAGTGTATTAAGACGCTGTGTTGTGATGTTAATGAT 3879
Qy 2563 A 2563
Db 3880 A 3880

RESULT 9

US-11-140-417-3
; Sequence 3, Application US/11140417
; Publication No. US20050266515A1
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Deborah A
; APPLICANT: Eddy, Edward M
; TITLE OF INVENTION: GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE-S (GAPDS), A GLYCO- TIC
; TITLE OF INVENTION: ENZYME EXPRESSED ONLY IN MALE GERM CELLS, IS A TARGET FOR M LE
; FILE REFERENCE: 421/76/2 PCT/CIP
; CURRENT APPLICATION NUMBER: US/11/140,417
; CURRENT FILING DATE: 2005-05-27
; PRIOR APPLICATION NUMBER: US 60/429,638
; PRIOR FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: PCT/US2003/037800
; PRIOR FILING DATE: 2003-11-26
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3
; LENGTH: 1450
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (4)..(1302)
US-11-140-417-3

Query Match 1.1%; Score 39.8; DB 7; Length 1450;
Best Local Similarity 51.4%; Pred. No. 1.2;
Matches 92; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

Qy 507 TGGTCTCTCCCAATTCGAATTATGCGGATTACACCGGCGAGCAAGCTCCGACCTAG 566
Db 107 TGATTCAGCACCTTCACCTCCACCCAGTTGAGAGCTTCGCGCACCCAGAGAGAGC 166
Qy 567 CCACAACACTCTCAACTATCTGCTCCGCGAGCAAAATGAGCCGACGCTTCTCTCCGA 626
Db 167 CGCGGCCACCGCCACACCGCTCTCTCAACCCAGATGAGCCAGAGAGCTTAAAGG 226
Qy 627 CTTCATTTACCTCCGATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 685
Db 227 CTCACCTCCACCTCCACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 285

RESULT 10

US-11-121-086-3/c
; Sequence 3, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138,6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04

; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3
; LENGTH: 168516
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-3

Query Match 1.1%; Score 39.8; DB 7; Length 168516;
Best Local Similarity 56.5%; Pred. No. 28;
Matches 74; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

Qy 556 CTCGACATGACCAACACCTCTCACTATCTGCTCGCGCAAGAGCCGACCGT 615
Db 68578 CTCCTACTCTCCCTTGGCTCCCTCCCTCCACCTCTTACTCTCCCTCTCTCCCT 68519
Qy 616 CTTCCTCGACATTCATTTCACTCGATTCCTCTCTCTCTCTCTCTCTCTCTCTCT 675
Db 68518 TCTCTCTCGCTCAACCTCTCTGCGCTTCCCTCCCTCTCTCTCTCTCTCTCTCT 68459
Qy 676 ACCACGCGCAC 686
Db 68458 CCTCTCTCTC 68448

RESULT 11

US-10-750-185-3575
; Sequence 3575, Application US/10750185
; Publication No. US2005026603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM11100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3575
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Bovine MMBT16645
US-10-750-185-3575

Query Match 1.1%; Score 39.6; DB 6; Length 600;
Best Local Similarity 50.0%; Pred. No. 0.74;
Matches 99; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

Qy 289 ATTTAATCATATAAGTCGTCGACGTGATACATATAGATTGTTTTTATTTTTC 348
Db 4 ATATCACTTATCAACACCTGAATACTAGCTAGCAAACTGTGAGATAATTAAT 63
Qy 349 GTACTGTGATGTTTTTTTATTTAATTTACTATCTCAAAATCAAAATTCATAACC 408
Db 64 ATGATGTATATATATATCAAGTTGAAGTTTACAGAGAAATAATTTTGGCAATATGA 123
Qy 409 TAGACGACCAAAAGTCTCTTCAATATGTAAACAGAACTTTTGTAGAGCTTA 468
Db 124 AATTAAGCAATTAATTCATTTTATACCTTAAGATATATAGCATGTGACTTATAT 183
Qy 469 AAAGACATCCCATGAA 486
Db 184 AAAAGATTCGTGGA 201

RESULT 12

; TYPE: DNA
; ORGANISM: Mus musculus
US-11-140-417-22

Query Match 1.1%; Score 39.2; DB 7; Length 11462;
Best Local Similarity 51.7%; Pred. No. 6.8;
Matches 89; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

OY	514	TCCCATTC	CAATT	TATG	CCGAT	TTAC	CCGCGG	AGAG	CAAA	AGCT	CCG	AGT	AGCC	CAAC	573
DB	3221	TCACAC	CCCG	CAAG	GTGA	AAAC	AGCA	CGCAC	CGCG	CGCC	CACT	CC	CTC	CAAC	3280
OY	574	ACCTT	ACA	CTAT	CTG	CTCC	GC	CA	GA	AA	TGG	CC	GA	CCG	CTT
DB	3281	ACCACT	CTCT	CTCT	CTCC	AG	AT	AG	CC	AG	AT	AG	CC	AG	AT
OY	634	TTCA	CTCG	ATTC	CTCT	CTCT	CTCT	CTCT	CTCT	CTCT	CTCT	CTCT	CTCT	CTCT	685
DB	3341	TC	CCCT	CCCT	CCCT	CCCT	CCCT	CCCT	CCCT	CCCT	CCCT	CCCT	CCCT	CCCT	3392

Search completed: December 10, 2005, 18:24:17
Job time : 325.571 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 10, 2005, 06:08:12 ; Search time 2797.5 Seconds
(without alignments)
10839.614 Million cell updates/sec

Title: US-10-600-070b-3
Perfect score: 3667
Sequence: 1 tgcctgcattcaagagagaat.....ctatacataagggctacaa 3667

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications_NA_Main:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3667	100.0	3667	US-10-600-070-3	Sequence 3, Appl1
2	3663.8	99.9	3667	US-10-600-070-10	Sequence 10, Appl1
3	1968.4	53.7	2679	US-10-739-930-227	Sequence 227, App
4	1931	52.7	2637	US-10-600-070-130	Sequence 130, App
5	1724	47.0	2406	US-10-600-070-1	Sequence 1, Appl1
6	1720.8	46.9	2406	US-10-600-070-9	Sequence 9, Appl1
7	1717.6	46.8	2406	US-10-600-070-128	Sequence 128, App
8	402.2	11.0	561	US-10-600-070-184	Sequence 184, App
9	295.6	8.1	631	US-10-600-070-184	Sequence 184, App
10	262.4	7.2	1146	US-10-424-599-119007	Sequence 12507, App
11	208.8	5.7	660	US-10-600-070-135	Sequence 135, App
12	194	5.3	2283	US-10-600-070-126	Sequence 126, App
13	178.6	4.9	607	US-10-021-323-3562	Sequence 3562, App
14	175	4.8	439	US-09-732-627A-2154	Sequence 2154, App
15	161	4.4	545	US-10-600-070-175	Sequence 175, App
16	159.4	4.3	552	US-10-021-323-3536	Sequence 3536, App
17	158	4.3	608	US-10-600-070-137	Sequence 137, App
18	149.4	4.1	1411	US-10-425-115-81853	Sequence 81853, A
19	146.2	4.0	2130	US-10-437-963-69932	Sequence 69932, A
20	145.6	4.0	527	US-10-600-070-134	Sequence 134, App
21	145.4	4.0	647	US-10-767-701-4069	Sequence 4069, App
22	141	3.8	1039	US-10-424-599-35059	Sequence 35059, A
23	139.8	3.8	1536	US-10-425-115-57452	Sequence 57452, A

24	136.2	3.7	647	US-10-600-070-185	Sequence 185, App
25	134.6	3.7	652	US-10-600-070-186	Sequence 186, App
26	133.2	3.6	563	US-10-600-070-178	Sequence 178, App
27	130.2	3.6	446	US-10-600-070-151	Sequence 151, App
28	127.6	3.5	479	US-10-600-070-150	Sequence 143, App
29	125	3.4	537	US-10-600-070-143	Sequence 138, App
30	122.6	3.3	307	US-10-600-070-138	Sequence 153, App
31	118.6	3.2	871	US-10-600-070-153	Sequence 69933, A
32	117.8	3.2	1703	US-10-437-963-69933	Sequence 8187, App
33	110.4	3.0	618	US-10-437-963-8187	Sequence 174, App
34	93.6	2.6	491	US-10-600-070-174	Sequence 136, App
35	92.6	2.5	187	US-10-600-070-136	Sequence 144, App
36	92	2.5	418	US-10-600-070-144	Sequence 176, App
37	91.4	2.5	460	US-10-600-070-187	Sequence 147, App
38	90.4	2.5	420	US-10-600-070-176	Sequence 148, App
39	89.4	2.4	604	US-10-600-070-147	Sequence 9474, App
40	89.4	2.4	653	US-10-600-070-148	Sequence 145, App
41	89.4	2.4	1032	US-10-767-701-9474	Sequence 180, App
42	87.2	2.4	480	US-10-600-070-145	Sequence 154, App
43	86.8	2.4	622	US-10-600-070-146	
44	85.2	2.3	300	US-10-600-070-180	
45	83	2.3	541	US-10-600-070-154	

ALIGNMENTS

RESULT 1
US-10-600-070-3
; Sequence 3, Application US/10600070
; Publication No. US20040139500A1
; GENERAL INFORMATION:
; APPLICANT: Osteoryoung, Katherine W.
; APPLICANT: Vilcha, Stanislav
; APPLICANT: Koksharova, Olga A.
; TITLE OF INVENTION: Plastid Division and Related Genes and Proteins, and Methods of
; FILE OF INVENTION: Use
; FILE REFERENCE: MSU-08153
; CURRENT APPLICATION NUMBER: US/10/600,070
; CURRENT FILING DATE: 2003-06-20
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 3667
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; US-10-600-070-3

Query Match 100.0%; Score 3667; DB 7; Length 3667;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3667; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	TGTTCTGCTTAAGAGAAATACATTTATAGCAATTTGCTTGAATTCACAGATTTTG	60
DB	1	TGTTCTGCTTAAGAGAAATACATTTATAGCAATTTGCTTGAATTCACAGATTTTG	60
QY	61	CTTGCTATAGGATTCATGCTGCTGCTTTGCTTACATTAATGATATAGTTTG	120
DB	61	CTTGCTATAGGATTCATGCTGCTGCTTTGCTTACATTAATGATATAGTTTG	120
QY	121	AATTTACATTTCACTTGATGTTAAGAAAGAGAGATTCAGGCTTTGTCGG	180
DB	121	AATTTACATTTCACTTGATGTTAAGAAAGAGAGATTCAGGCTTTGTCGG	180
QY	181	TTTAACTTTAAATGATCAAGATTAAGTATGCTTACTGCTCTATATGTTAA	240
DB	181	TTTAACTTTAAATGATCAAGATTAAGTATGCTTACTGCTCTATATGTTAA	240
QY	241	AATAAGCACTCCACAGGTTCTTAGTGGAATAGATTATTAAGAGATTTACATCAT	300
DB	241	AATAAGCACTCCACAGGTTCTTAGTGGAATAGATTATTAAGAGATTTACATCAT	300

```
Qy 301 AAAGTCCGTGCGAGCTGTAACTCATAGATTGTTTTTATTTTTTTCAGTAGCTGTGAT 360
Db 301 AAAGTCCGTGCGAGCTGTAACTCATAGATTGTTTTTATTTTTTTCAGTAGCTGTGAT 360
Qy 361 GTTTTTGATTTAACTTAACTACTCAAAATCAAAATCCATPAAACCTTAGACGACCAA 420
Db 361 GTTTTTGATTTAACTTAACTACTCAAAATCAAAATCCATPAAACCTTAGACGACCAA 420
Qy 421 CAGTCTCTTCAATATGTAAAAACAGAACAAAGTTTTTGTAGTAGCCTTAAAAAGACTCC 480
Db 421 CAGTCTCTTCAATATGTAAAAACAGAACAAAGTTTTTGTAGTAGCCTTAAAAAGACTCC 480
Qy 481 ATGGAAGCTCTGAGTCACTGCGGCTTGTGTCTCTCCCATTCCAATTATGCCATTACCA 540
Db 481 ATGGAAGCTCTGAGTCACTGCGGCTTGTGTCTCTCCCATTCCAATTATGCCATTACCA 540
Qy 541 CCGGCGAGCAAAAGCTCGAGCTAGCACAACCTCTACAACTATGTGCTCGCGCAGC 600
Db 541 CCGGCGAGCAAAAGCTCGAGCTAGCACAACCTCTACAACTATGTGCTCGCGCAGC 600
Qy 601 AAATGGGCGGACCGGTCTTCTCCGACTTCAATTTCACTCGGATTCCTCCCTCTCC 660
Db 601 AAATGGGCGGACCGGTCTTCTCCGACTTCAATTTCACTCGGATTCCTCCCTCTCC 660
Qy 661 TTCCGCAACCGGCAACACACACGCGCACTCTCGTCTCTGCGCAACCATCTATGTATGCC 720
Db 661 TTCCGCAACCGGCAACACACACGCGCACTCTCGTCTCTGCGCAACCATCTATGTATGCC 720
Qy 721 GAAAGCCAGCTCCCATCCCATTTGATTCTACAGAGTATTAGGAGCTCAAAACATTTG 780
Db 721 GAAAGCCAGCTCCCATCCCATTTGATTCTACAGAGTATTAGGAGCTCAAAACATTTG 780
Qy 781 TTAACCGATGAATCAGAAAGAGCATTCGAGCTAGGGTTTGGAAACCGCGCAATTCGGT 840
Db 781 TTAACCGATGAATCAGAAAGAGCATTCGAGCTAGGGTTTGGAAACCGCGCAATTCGGT 840
Qy 841 TTCAGCGAGCGAGCTTTAATCAGCCGAGACAGATTTCAAGCTGTGGGAAACTGTG 900
Db 841 TTCAGCGAGCGAGCTTTAATCAGCCGAGACAGATTTCAAGCTGTGGGAAACTGTG 900
Qy 901 TCTAATCTCGGCTTGAAGAGAGTACAAATGAAGGCTTCTTGAATGAGAAGACTACA 960
Db 901 TCTAATCTCGGCTTGAAGAGAGTACAAATGAAGGCTTCTTGAATGAGAAGACTACA 960
Qy 961 GTCATCATCTGATGTTCTTTGGGATPAGGTAAATTCGATTTGCGAAATATAAAGTTTCTC 1020
Db 961 GTCATCATCTGATGTTCTTTGGGATPAGGTAAATTCGATTTGCGAAATATAAAGTTTCTC 1020
Qy 1021 GTTTTAAATTTGATGAATTTGAATGAAGAAAGAACTTTATCTAGTGAAGGTTCTCGGCGC 1080
Db 1021 GTTTTAAATTTGATGAATTTGAATGAAGAAAGAACTTTATCTAGTGAAGGTTCTCGGCGC 1080
Qy 1081 TCTCTGTGTATTGCAAGAGGTGTGAGACTGAGATAGTCTTCTCGGGTGTGTAGGCTCT 1140
Db 1081 TCTCTGTGTATTGCAAGAGGTGTGAGACTGAGATAGTCTTCTCGGGTGTGTAGGCTCT 1140
Qy 1141 GCTTAAAGAGAGGTTGCTTAAAGTCGTTTAAACAAAGATGTGTTTATGTTATGCGCTTGC 1200
Db 1141 GCTTAAAGAGAGGTTGCTTAAAGTCGTTTAAACAAAGATGTGTTTATGTTATGCGCTTGC 1200
Qy 1201 GTTTCCTGATGTCTGAGAGGATGCTATGCACTTGATCACTGATTTTATTACTGGTTA 1260
Db 1201 GTTTCCTGATGTCTGAGAGGATGCTATGCACTTGATCACTGATTTTATTACTGGTTA 1260
Qy 1261 TGAAGTTGTGAGGAAGCTTTGAAGCTTTTACAGGTATGTTGACTTGTGGTAAATTTG 1320
Db 1261 TGAAGTTGTGAGGAAGCTTTGAAGCTTTTACAGGTATGTTGACTTGTGGTAAATTTG 1320
Qy 1321 ACGAGCGTTGCTTTAATAAGAACTTCTTGAATTTGATACTTTGTTATTTAGACTTGTGTA 1380
Db 1321 ACGAGCGTTGCTTTAATAAGAACTTCTTGAATTTGATACTTTGTTATTTAGACTTGTGTA 1380
Qy 1381 GGAGGAAGGAGCAAGTACCTTGCACCGGATTTACGTGCAAAATGTATGATGACTTTGGA 1440
Db 1381 GGAGGAAGGAGCAAGTACCTTGCACCGGATTTACGTGCAAAATGTATGATGACTTTGGA 1440
Qy 1441 AGAGTCACTCCGCGTATGCTTTGAGACTACTTGGCTTACCGCTGTGATGATTAACG 1500
Db 1441 AGAGTCACTCCGCGTATGCTTTGAGACTACTTGGCTTACCGCTGTGATGATTAACG 1500
Qy 1501 TCGAAAAAGACTAATGTTTTAAGCGGTGTGCGGAATATTTTGTGTCTGTGTGAGGAGG 1560
Db 1501 TCGAAAAAGACTAATGTTTTAAGCGGTGTGCGGAATATTTTGTGTGTGTGTGAGGAGG 1560
Qy 1561 TGGAGCATCAGCTCTTGTGGGGGTTTGAACCGGTGAAGAAATTTATGATAGAGGCTTTT 1620
Db 1561 TGGAGCATCAGCTCTTGTGGGGGTTTGAACCGGTGAAGAAATTTATGATAGAGGCTTTT 1620
Qy 1621 ACGAATACAGCTCTGAGCAGGTATACAGTTTATGATACCTTTTAAATTTCTTTAGCA 1680
Db 1621 ACGAATACAGCTCTGAGCAGGTATACAGTTTATGATACCTTTTAAATTTCTTTAGCA 1680
Qy 1681 TGAATATACTTAAAGTTTCTCATTTTAAATGATATGTGTGTAGTGAATCTTTTGTGA 1740
Db 1681 TGAATATACTTAAAGTTTCTCATTTTAAATGATATGTGTGTAGTGAATCTTTTGTGA 1740
Qy 1741 GCTAACCCCAAGCAATATTTCCAGCAGAGTCAATTGAAGTTTACGAAGTTGCACTTGCCTT 1800
Db 1741 GCTAACCCCAAGCAATATTTCCAGCAGAGTCAATTGAAGTTTACGAAGTTGCACTTGCCTT 1800
Qy 1801 GTGGCTCAAGCTTTTATTTGTTAAGAAAGCACAACCTTTTACAGATGCTGATTAAGCAATTC 1860
Db 1801 GTGGCTCAAGCTTTTATTTGTTAAGAAAGCACAACCTTTTACAGATGCTGATTAAGCAATTC 1860
Qy 1861 CAGCACTTACAGCAGGCTAATGATATGCTATGAGATTCCTCGATGTTGTATGATACA 1920
Db 1861 CAGCACTTACAGCAGGCTAATGATATGCTATGAGATTCCTCGATGTTGTATGATACA 1920
Qy 1921 CGGAATATTTGGAGATGAGCTTCCGTAGAAAGGGGACTCTGTGACGCTTATAGGC 1980
Db 1921 CGGAATATTTGGAGATGAGCTTCCGTAGAAAGGGGACTCTGTGACGCTTATAGGC 1980
Qy 1981 AAAGTTGATGAATGCTGATATGTTGGGCTTGAACAGTGAAGATTCACAATATAGCAAT 2040
Db 1981 AAAGTTGATGAATGCTGATATGTTGGGCTTGAACAGTGAAGATTCACAATATAGCAAT 2040
Qy 2041 CCAAGCTATTTGAGATTTGTTTGGAGATTCAAATGATGATGATGATGATGATGATGATGAT 2100
Db 2041 CCAAGCTATTTGAGATTTGTTTGGAGATTCAAATGATGATGATGATGATGATGATGATGAT 2100
Qy 2101 GGAATATGCAAAATTTGTTGAAACCTGTTGGCAGGGGTTTCTTCTAGGTTCAAGAC 2160
Db 2101 GGAATATGCAAAATTTGTTGAAACCTGTTGGCAGGGGTTTCTTCTAGGTTCAAGAC 2160
Qy 2161 ACCAATGATAAAAATTTTAACTCGGGGACTACATATGATGATGATGATGATGATGATGATGAT 2220
Db 2161 ACCAATGATAAAAATTTTAACTCGGGGACTACATATGATGATGATGATGATGATGATGATGAT 2220
Qy 2221 TTGGAAGAGATGAGGATGTTCAAGGTTTCTCTTATGACTCTGCACTATGCGCAAG 2280
Db 2221 TTGGAAGAGATGAGGATGTTCAAGGTTTCTCTTATGACTCTGCACTATGCGCAAG 2280
Qy 2281 ATTGAGCCGAGCATGTGAAGAGTATGCTATGCAAGGCACTGCAAGAAAGTTTCTCTCC 2340
Db 2281 ATTGAGCCGAGCATGTGAAGAGTATGCTATGCAAGGCACTGCAAGAAAGTTTCTCTCC 2340
Qy 2341 CGCTATACATATGAATACTCGGCTGAACCCAAAGATGTGCAAGAGACAGTGTATGATGTA 2400
Db 2341 CGCTATACATATGAATACTCGGCTGAACCCAAAGATGTGCAAGAGACAGTGTATGATGTA 2400
Qy 2401 GATCTGTTGTGTAACAATGTAGAGCGGTGATGTGAGCTGTGTCTTTATTTGCAAGACT 2460
Db 2401 GATCTGTTGTGTAACAATGTAGAGCGGTGATGTGAGCTGTGTCTTTATTTGCAAGACT 2460
Qy 2461 GTAAGACCTCTGAAAACTTTGAACTTAATGATTAAGCAATTTGAGAGTGGGGTCTGAG 2520
Db 2461 GTAAGACCTCTGAAAACTTTGAACTTAATGATTAAGCAATTTGAGAGTGGGGTCTGAG 2520
```

Db 2461 GTAAAGCCCTCTGAAAATTGAACTAAATGATATGCAATTCGAGCTGGGCTCTCAGAG 2520
Qy 2521 AGTAGCGTGTAGTAACTACTGTTGAAATGTCGGTGTGATATGTTAAAGAGCAAGT 2580
Db 2521 ACTAGGGTGTAGTAACTACTGTTGAAATGTCGGTGTGATATGTTAAAGAGCAAGT 2580
Qy 2581 GTGAAGATCTAGGCTGCTGGTGGCAATGGAATGCACTGTTCAAGCCAGAAATAT 2640
Db 2581 GTGAAGATCTAGGCTGCTGGTGGCAATGGAATGCACTGTTCAAGCCAGAAATAT 2640
Qy 2641 TTTCTTAAAGAGCTCATCTTTTCAAGCAGAGATATGTTCTCTATGAAATCTGAT 2700
Db 2641 TTTCTTAAAGAGCTCATCTTTTCAAGCAGAGATATGTTCTCTATGAAATCTGAT 2700
Qy 2701 GTGCGTACCATAGTATGATTAATGATGCAATTTTCTATATCTGATTCCTCAAAATA 2760
Db 2701 GTGCGTACCATAGTATGATTAATGATGCAATTTTCTATATCTGATTCCTCAAAATA 2760
Qy 2761 TGCTTGTGTTGTAGCTAAGAAATATGTTCCCATTAATATGTTCCCAAAAGTTGTACC 2820
Db 2761 TGCTTGTGTTGTAGCTAAGAAATATGTTCCCATTAATATGTTCCCAAAAGTTGTACC 2820
Qy 2821 AAGATTAACAAGTGTGCTGTAATTTTCACTAATATGCTGTAATTTTGTATCAAA 2880
Db 2821 AAGATTAACAAGTGTGCTGTAATTTTCACTAATATGCTGTAATTTTGTATCAAA 2880
Qy 2881 CTGTAGACAGAAATGTAAATTTTCACTCAACATTTCTGTTAGAAATAAGTATGATAG 2940
Db 2881 CTGTAGACAGAAATGTAAATTTTCACTCAACATTTCTGTTAGAAATAAGTATGATAG 2940
Qy 2941 AGATTGCTTATGTTGTGGCTTTGTCCAATTTTCTTTCCTTGAATTTTCTTTTGCATTT 3000
Db 2941 AGATTGCTTATGTTGTGGCTTTGTCCAATTTTCTTTCCTTGAATTTTCTTTTGCATTT 3000
Qy 3001 AGGCTACGTCAGAGCTGACGATTCAGAAAGCACTTCCAGAAATGATGCTAGAGATGCGAGA 3060
Db 3001 AGGCTACGTCAGAGCTGACGATTCAGAAAGCACTTCCAGAAATGATGCTAGAGATGCGAGA 3060
Qy 3061 GAATATAGTATCCAAAGTGCAGAAATTAAGTCTCTGCTTTTGGGCTGATCAACGCAAT 3120
Db 3061 GAATATAGTATCCAAAGTGCAGAAATTAAGTCTCTGCTTTTGGGCTGATCAACGCAAT 3120
Qy 3121 AGAAATGTTTACCAAGAGTGGAGGAATTAATCTACATTCATCAATTTGTGTAATAAATCTGT 3180
Db 3121 AGAAATGTTTACCAAGAGTGGAGGAATTAATCTACATTCATCAATTTGTGTAATAAATCTGT 3180
Qy 3181 TGGACATGATTAATCTGCTGCTGCTGTTGATTTCTGTTATTAATGATTTTGGATGGGC 3240
Db 3181 TGGACATGATTAATCTGCTGCTGCTGTTGATTTCTGTTATTAATGATTTTGGATGGGC 3240
Qy 3241 GAATGCTGAAGATTTGGAAGTGCAGAGCAAGTGAATGCGAGGCTGGGTTGTTTATG 3300
Db 3241 GAATGCTGAAGATTTGGAAGTGCAGAGCAAGTGAATGCGAGGCTGGGTTGTTTATG 3300
Qy 3301 ATTATATACCTGTTGAACTATCTGTTGACAGTGTGACAGTCTCAGCAAGTGAACCCGCG 3360
Db 3301 ATTATATACCTGTTGAACTATCTGTTGACAGTGTGACAGTCTCAGCAAGTGAACCCGCG 3360
Qy 3361 CTCTGCTGGAAGCACTCTGAGAGAGTCTGCTGTTCTATCTGATTTTGGTTCTCCAGAAA 3420
Db 3361 CTCTGCTGGAAGCACTCTGAGAGAGTCTGCTGTTCTATCTGATTTTGGTTCTCCAGAAA 3420
Qy 3421 ACAATGCTACTGATGTCAGAACTTACACAAACAAAGATTTTCTGCTCAAGTCAAG 3480
Db 3421 ACAATGCTACTGATGTCAGAACTTACACAAACAAAGATTTTCTGCTCAAGTCAAG 3480
Qy 3481 GGTGGAATAATCACTGAAGGCTCTGTTCTTGCAATCAATAATATGATGCAATGCTCT 3540
Db 3481 GGTGGAATAATCACTGAAGGCTCTGTTCTTGCAATCAATAATATGATGCAATGCTCT 3540
Qy 3541 GAGCTTGGGAAGATCTCTTGTGTTCTGTAATTTCTCTCTCAAGTGAATGTTAAATGA 3600
Db 3541 GAGCTTGGGAAGATCTCTTGTGTTCTGTAATTTCTCTCTCAAGTGAATGTTAAATGA 3600

Qy 3601 ACACAAAAAATTAAAGTCTTGTGGCAGACCCCTTTCTCTGATCTAATATTAACATAAG 3660
Db 3601 ACACAAAAAATTAAAGTCTTGTGGCAGACCCCTTTCTCTGATCTAATATTAACATAAG 3660
Qy 3661 GCTACAA 3667
Db 3661 GCTACAA 3667

RESULT 2
US-10-600-070-10
; Sequence 10, Application US/10600070
; Publication No. US20040139500A1
; GENERAL INFORMATION:
; APPLICANT: Oesteryoung, Katherine W.
; APPLICANT: Vittha, Stanislaw
; APPLICANT: Koksharova, Olga A.
; TITLE OF INVENTION: Plasticid Division and Related Genes and Proteins, and Methods of
; FILE REFERENCE: MSU-08153
; CURRENT APPLICATION NUMBER: US/10/600,070
; CURRENT FILING DATE: 2003-06-20
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 10
; LENGTH: 3667
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-600-070-10

Query Match 99.9%; Score 3663.8; DB 7; Length 3667;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3665; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TGTTCGATTAAAGAGATACAAATTATAGCAATTTGCTGATTTCAACAGATTTTG 60
Db 1 TGTTCGATTAAAGAGATACAAATTATAGCAATTTGCTGATTTCAACAGATTTTG 60
Qy 61 CTGGCTATAGATTCATTTGCTGCTGTTTCTTTTACATTTACATGTCATATAGTTTCG 120
Db 61 CTGGCTATAGATTCATTTGCTGCTGTTTCTTTTACATTTACATGTCATATAGTTTCG 120
Qy 121 AATTTTACATTTTCAAGTGTGATTTAAGAAAGAGAGGAATGATGGGTTTGTGGG 180
Db 121 AATTTTACATTTTCAAGTGTGATTTAAGAAAGAGAGGAATGATGGGTTTGTGGG 180
Qy 181 TTTTAACTTTTAAAGTACAAAGATTAAGTCAATGTTTACTGTTGCTCTATATGTGTA 240
Db 181 TTTTAACTTTTAAAGTACAAAGATTAAGTCAATGTTTACTGTTGCTCTATATGTGTA 240
Qy 241 AATGAGGCACTCCACGCTTCTTAGGTGAATATGATTTATGAGCATTTTACATCAT 300
Db 241 AATGAGGCACTCCACGCTTCTTAGGTGAATATGATTTATGAGCATTTTACATCAT 300
Qy 301 AAAGTCGTCGCACTGTAATCATATGATTTTATTTTATTTTTCAGTAGCTGTGAT 360
Db 301 AAAGTCGTCGCACTGTAATCATATGATTTTATTTTATTTTTCAGTAGCTGTGAT 360
Qy 361 GTTTTGTGATTTTAACTTATATCTCAAAATCAAAATTCATAAACCTTAGACGACAAA 420
Db 361 GTTTTGTGATTTTAACTTATATCTTACGCAAAATCAAAATTCATAAACCTTAGACGACAAA 420
Qy 421 CAGTCTCTTCAATATGTAACAGAACAAAGTTTGTAGTATGCTTAAAGACACTGCC 480
Db 421 CAGTCTCTTCAATATGTAACAGAACAAAGTTTGTAGTATGCTTAAAGACACTGCC 480
Qy 481 ATGGAAGCTGAGTCAAGTGGGATTTGCTCTCCCATTTCCAAATATAGCCGATTAACA 540
Db 481 ATGGAAGCTGAGTCAAGTGGGATTTGCTCTCCCATTTCCAAATATAGCCGATTAACA 540
Qy 541 CCGGCGACGACAAAGTCCGAGTGAAGCCACAAACCTCTACAACTATCTGCTCCGACG 600

```
|||||
Db 541 CCGGCGACGACAAAGCTCCGACGTAGCCACAACCTCTACAACTATCTGCTCCGCGACG 600
Qy AAATGGGCGCGACCGCTCTCTCCGACTTCAATTTCACTCCGATTCCTCTCCCTCC 660
Db 601 AAATGGGCGCGACCGCTCTCTCCGACTTCAATTTCACTCCGATTCCTCTCCCTCC 660
Qy 601 AAATGGGCGCGACCGCTCTCTCCGACTTCAATTTCACTCCGATTCCTCTCCCTCC 660
Db 661 TTGCGCACCGCGACCGCGACCGCGACCTCTCGTCTCTCCGACGATCTTAATGATGCTCC 720
Qy 661 TTGCGCACCGCGACCGCGACCGCGACCTCTCGTCTCTCCGACGATCTTAATGATGCTCC 720
Db 661 TTGCGCACCGCGACCGCGACCGCGACCTCTCGTCTCTCCGACGATCTTAATGATGCTCC 720
Qy 721 GAAGCGCACGCTCCCATCTCCCATTTGATTTCTACGAGGATTAAGAGCTCAACACATTC 780
Db 721 GAAGCGCACGCTCCCATCTCCCATTTGATTTCTACGAGGATTAAGAGCTCAACACATTC 780
Qy 721 GAAGCGCACGCTCCCATCTCCCATTTGATTTCTACGAGGATTAAGAGCTCAACACATTC 780
Db 721 GAAGCGCACGCTCCCATCTCCCATTTGATTTCTACGAGGATTAAGAGCTCAACACATTC 780
Qy 781 TTAACCGATGGAAATCAGAAAGACATTCGAGGTAGGGTTTGAAACCGCGCAATTCGAT 840
Db 781 TTAACCGATGGAAATCAGAAAGACATTCGAGGTAGGGTTTGAAACCGCGCAATTCGAT 840
Qy 841 TTCAAGCGACGACGCTTTAATCAGCCGAGACAGATTCCTCAAGCTGCTTGCGAAACTCTG 900
Db 841 TTCAAGCGACGACGCTTTAATCAGCCGAGACAGATTCCTCAAGCTGCTTGCGAAACTCTG 900
Qy 901 TCTAATCTCGGCTCTAGAGAGAGATCAATGAGGTCTTCTGATGATGAAGAACTTCA 960
Db 901 TCTAATCTCGGCTCTAGAGAGAGATCAATGAGGTCTTCTGATGATGAAGAACTTCA 960
Qy 961 GTCAATCACTGATGTTCTTGGGATTAAGGTAAATTCGATTTGCGAATTAATAAGTTCTTC 1020
Db 961 GTCAATCACTGATGTTCTTGGGATTAAGGTAAATTCGATTTGCGAATTAATAAGTTCTTC 1020
Qy 1021 GTTTAATTTTATGATGATTAAGAAAGAACTTTATCTAGTGAAGTTCCTGCGGC 1080
Db 1021 GTTTAATTTTATGATGATTAAGAAAGAACTTTATCTAGTGAAGTTCCTGCGGC 1080
Qy 1081 TCTCTGTGTATTTGCAAGAGGTGTGAGACTGAGATAGTTCTTGCGGTTGTGAGGCTCT 1140
Db 1081 TCTCTGTGTATTTGCAAGAGGTGTGAGACTGAGATAGTTCTTGCGGTTGTGAGGCTCT 1140
Qy 1141 GCTTAAGGAGAGGTGCTTAAGCTTAAGCTTAAGAGATGTTGTTAGTTAGGCGCTTC 1200
Db 1141 GCTTAAGGAGAGGTGCTTAAGCTTAAGCTTAAGAGATGTTGTTAGTTAGGCGCTTC 1200
Qy 1201 GTTCTCGATGTCTCGAGGGAATGCTATGCGAATTCGATTCACCTGATTTTATTAATGCTGTTA 1260
Db 1201 GTTCTCGATGTCTCGAGGGAATGCTATGCGAATTCGATTCACCTGATTTTATTAATGCTGTTA 1260
Qy 1261 TGAATTTGTGAGAGAGCTTTGAAGCTTTTACAGGTAGTTGACTTGCTTTGGTAAATTTG 1320
Db 1261 TGAATTTGTGAGAGAGCTTTGAAGCTTTTACAGGTAGTTGACTTGCTTTGGTAAATTTG 1320
Qy 1321 ACGAGCGTTGCTTAATAAGAACTTCTGATTTGATAGCTTTGATTTAGCTTTGTTA 1380
Db 1321 ACGAGCGTTGCTTAATAAGAACTTCTGATTTGATAGCTTTGATTTAGCTTTGTTA 1380
Qy 1381 GGAAGGAGAGACAGTACCTTGCACCGGATTTTACGTGCACAAAATGATGAAGACTTTGGA 1440
Db 1381 GGAAGGAGAGACAGTACCTTGCACCGGATTTTACGTGCACAAAATGATGAAGACTTTGGA 1440
Qy 1441 AGAGATCACTCGCGCTTATGCTTTGAGAGCTTACCTTGCTTACCGCTTGCTGATTAAGC 1500
Db 1441 AGAGATCACTCGCGCTTATGCTTTGAGAGCTTACCTTGCTTACCGCTTGCTGATTAAGC 1500
Qy 1501 TGGGAAAAGACCTAAATGCTTAAAGCGGTGCGGAAATTTTGTGCTGTTGAGAGAG 1560
Db 1501 TGGGAAAAGACCTAAATGCTTAAAGCGGTGCGGAAATTTTGTGCTGTTGAGAGAG 1560
Qy 1561 TGGAGCATCAGCTCTTGTGGGGGTTTGACCGGTGAGAACTTTATGATGAGGGCTTTT 1620
Db 1561 TGGAGCATCAGCTCTTGTGGGGGTTTGACCGGTGAGAACTTTATGATGAGGGCTTTT 1620
Qy 1621 ACGAATGACAGCTGCTGAGAGATTAACGTTTATGATCTTTTATTTTATTTTACGA 1680
Db 1621 ACGAATGACAGCTGCTGAGAGATTAACGTTTATGATCTTTTATTTTATTTTACGA 1680

|||||
Db 1621 ATGAATGACAGCTGCTGAGAGATTAACGTTTATGATCTTTTATTTTATTTTACGA 1680
Qy 1681 TGAATATATTTAGGTTTCTCATTTTATATGATGTTGTGATGAGTTGATCTTTTGTGA 1740
Db 1681 TGAATATATTTAGGTTTCTCATTTTATATGATGTTGTGATGAGTTGATCTTTTGTGA 1740
Qy 1741 GCTAACCCAGCAATATTCAGAGAGTCAATTTGAAGTTTGAAGTTGACTTTCCTT 1800
Db 1741 GCTAACCCAGCAATATTCAGAGAGTCAATTTGAAGTTTGAAGTTGACTTTCCTT 1800
Qy 1801 GTGCTCAAGCTTTTATGTTGAAGAGCGACCTTTTACAGGATGCTGATTAAGCAATTC 1860
Db 1801 GTGCTCAAGCTTTTATGTTGAAGAGCGACCTTTTACAGGATGCTGATTAAGCAATTC 1860
Qy 1861 CAGCAACTTCAGAGAGCTAAGTATAGCTATGAGATTCCTGATGTTGATGATGATCA 1920
Db 1861 CAGCAACTTCAGAGAGCTAAGTATAGCTATGAGATTCCTGATGTTGATGATGATCA 1920
Qy 1921 CCGAATATATGAGAGATGACTTCGCTTGAAGAGGAGACTCTGTGCACTGTTATAGGC 1980
Db 1921 CCGAATATATGAGAGATGACTTCGCTTGAAGAGGAGACTCTGTGCACTGTTATAGGC 1980
Qy 1981 AAGTTGATGATGATGCGGTATGCTTGGGCTTGAAGAGGAGATTCACAAATTAAGGAT 2040
Db 1981 AAGTTGATGATGATGCGGTATGCTTGGGCTTGAAGAGGAGATTCACAAATTAAGGAT 2040
Qy 2041 CCAGCTATTTGAGAGTTGTTTGAAGATTCAAATCGTATGACATGATGATCTTCCT 2100
Db 2041 CCAGCTATTTGAGAGTTGTTTGAAGATTCAAATCGTATGACATGATGATCTTCCT 2100
Qy 2101 GGAATATGAAATTTGTTGGAACCTGCTGAGAGGAGGTTCTTCTCTAGTTTCAAGAC 2160
Db 2101 GGAATATGAAATTTGTTGGAACCTGCTGAGAGGAGGTTCTTCTCTAGTTTCAAGAC 2160
Qy 2161 ACCAAAGATTAATAAATTTTAACTCGGGAGTACTATGATGATGATGATGATGATGATG 2220
Db 2161 ACCAAAGATTAATAAATTTTAACTCGGGAGTACTATGATGATGATGATGATGATGATG 2220
Qy 2221 TTGGAAGAGATGAGAGTGTGAGAGGTTCTCTTATGAGCTGCTGCAATATGAGAGG 2280
Db 2221 TTGGAAGAGATGAGAGTGTGAGAGGTTCTCTTATGAGCTGCTGCAATATGAGAGG 2280
Qy 2281 ATTGAGCCGAGAGATGGAAGCTAGTGTATGAGAGCTGCAAGAAAGTTTCTCTCC 2340
Db 2281 ATTGAGCCGAGAGATGGAAGCTAGTGTATGAGAGCTGCAAGAAAGTTTCTCTCC 2340
Qy 2341 CGCTATACAGATGAAACTCGGCTGAACCCAGAGATGTCAGAGAAAGTTTCTCTCC 2400
Db 2341 CGCTATACAGATGAAACTCGGCTGAACCCAGAGATGTCAGAGAAAGTTTCTCTCC 2400
Qy 2401 GATCTGTTGTAACAATGATGAGCCGTGATGAGAGCTGATGAGAGCTTATTTGAGAGACT 2460
Db 2401 GATCTGTTGTAACAATGATGAGCCGTGATGAGAGCTGATGAGAGCTTATTTATTTGAGAGACT 2460
Qy 2461 GTAAAGCCCTCTGAAAACTTTGAAACTTAATGATTAATGAAATTTGAGAGCTGAGAG 2520
Db 2461 GTAAAGCCCTCTGAAAACTTTGAAACTTAATGATTAATGAAATTTGAGAGCTGAGAG 2520
Qy 2521 AGTAGCGTTGATGAATACTAGTTGAAGATGCTGCTGATGATGATTAAGAGAGCAAGT 2580
Db 2521 AGTAGCGTTGATGAATACTAGTTGAAGATGCTGCTGATGATGATTAAGAGAGCAAGT 2580
Qy 2581 GTGAAGATCTTACGCTGCTGCTGAGAAATTTGAGCTGATTTCACTGTTCAAGCAGAGAT 2640
Db 2581 GTGAAGATCTTACGCTGCTGCTGAGAAATTTGAGCTGATTTCACTGTTCAAGCAGAGAT 2640
Qy 2641 TTTCTTAAAGAGCTCATCTTTTCAACGCAAGATATGTTTCTTCTATGATGATCTGAT 2700
Db 2641 TTTCTTAAAGAGCTCATCTTTTCAACGCAAGATATGTTTCTTCTATGATGATCTGAT 2700
Qy 2701 GTGCTATCAGATGATGATTAATGATGATGATTTTCAATATCTGATGCTCAAAATA 2760
Db 2701 GTGCTATCAGATGATGATTAATGATGATGATTTTCAATATCTGATGCTCAAAATA 2760
```

QY 2761 TGTCTGTTTGTGAGCTAAGACATAGTTCCTTAATACATGTCCTCCAAAAGTTGAC 2820
Db 2761 TGTCTGTTTGTGAGCTAAGACATAGTTCCTTAATACATGTCCTCCAAAAGTTGAC 2820
QY 2821 AAGATTAACAAGTGTCTGAGTAAATTTCACTAATATAGTGTGTAATTTTGTATCAA 2880
Db 2821 AAGATTAACAAGTGTCTGAGTAAATTTCACTAATATAGTGTGTAATTTTGTATCAA 2880
QY 2881 CTGTGACAGAAATTAATTTCACTCAACATTTCTGTGTAAGTAATAGTAAGATTAG 2940
Db 2881 CTGTGACAGAAATTAATTTCACTCAACATTTCTGTGTAAGTAATAGTAAGATTAG 2940
QY 2941 AGATTGCTTGAAGTGTGCTTGTGCTCAACTTTTCTTCTGTAATTTTCTTCTGATT 3000
Db 2941 AGATTGCTTGAAGTGTGCTTGTGCTCAACTTTTCTTCTGTAATTTTCTTCTGATT 3000
QY 3001 AGGTCATCAGAGCTGACATGACATGACATCTTCCCAAGTAATGATGCTAGACATGACGA 3060
Db 3001 AGGTCATCAGAGCTGACATGACATGACATCTTCCCAAGTAATGATGCTAGACATGACGA 3060
QY 3061 GAATATAGTATCAAGTGCAGAAAGATTAAAGTCTGAGCTTTTGGGCTGATCACCGCAT 3120
Db 3061 GAATATAGTATCAAGTGCAGAAAGATTAAAGTCTGAGCTTTTGGGCTGATCACCGCAT 3120
QY 3121 AGAAATGTACAGAGGTGAGGGAATTAATCTCAATCAATGTTGTGAAAACTGT 3180
Db 3121 AGAAATGTACAGAGGTGAGGGAATTAATCTCAATCAATGTTGTGAAAACTGT 3180
QY 3181 TGACATATTAATAGTGTGCTGCTGTTGTTGATTTCTGTTATTTAATGTTTGGATGGCC 3240
Db 3181 TGACATATTAATAGTGTGCTGCTGTTGTTGATTTCTGTTATTTAATGTTTGGATGGCC 3240
QY 3241 GAATGCTGAAGATTGGACTGACAGAGAGCTGAAACTGCGCAGCTTGGGTTGTTTATG 3300
Db 3241 GAATGCTGAAGATTGGACTGACAGAGAGCTGAAACTGCGCAGCTTGGGTTGTTTATG 3300
QY 3301 ATTATACACTGTGTAACATATCTGTTGACAGTGTGACAGTCTCAGCAGATGGAAACCGTG 3360
Db 3301 ATTATACACTGTGTAACATATCTGTTGACAGTGTGACAGTCTCAGCAGATGGAAACCGTG 3360
QY 3361 CTGTGTGGAAGCACTGTGAGAGAGTCTGTTCTGTAATGATTTGTTGTTATCCAGAA 3420
Db 3361 CTGTGTGGAAGCACTGTGAGAGAGTCTGTTCTGTAATGATTTGTTGTTATCCAGAA 3420
QY 3421 ACAATGCTACGATGTGCAAGCTTACACAGAAAGTAAGATTCTGTCGTCAGTACG 3480
Db 3421 ACAATGCTACGATGTGCAAGCTTACACAGAAAGTAAGATTCTGTCGTCAGTACG 3480
QY 3481 GGTGGAATACATGAAAGCTCTGTTCTTGATCATATAATATGATGATGATGATGATG 3540
Db 3481 GGTGGAATACATGAAAGCTCTGTTCTTGATCATATAATATGATGATGATGATGATG 3540
QY 3541 GAGCTTGGAGATTTCTTGTGTTGTAATTTCTCTCTAAGTTAGTTTAAATGA 3600
Db 3541 GAGCTTGGAGATTTCTTGTGTTGTAATTTCTCTCTAAGTTAGTTTAAATGA 3600
QY 3601 ACACAAAAAATTAAGTTTGTGACACCCCTTTCTTGAATCTAATTAACATAAG 3660
Db 3601 ACACAAAAAATTAAGTTTGTGACACCCCTTTCTTGAATCTAATTAACATAAG 3660
QY 3661 GCTACAA 3667
Db 3661 GCTACAA 3667

RESULT 3

US-10-739-930-227

; Sequence 227, Application US/10739930

; Publication No. US20040216190A1

; GENERAL INFORMATION:

; APPLICANT: Kovalic, David K.

; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH

; TITLE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53377)B
; CURRENT APPLICATION NUMBER: US/10/739,930
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 227
; LENGTH: 2679
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURES:
; OTHER INFORMATION: Clone ID: ARATH-23APR03-CLUSTER13643_1
US-10-739-930-227

Query Match 53.7%; Score 1968.4; DB 8; Length 2679;
Best Local Similarity 80.5%; Pred. No. 0;
Matches 2657; Conservative 0; Mismatches 11; Indels 632; Gaps 5;

QY 368 GATTATATTTACTACTCAAAATCAAAATCCCTAGACGACCAACGATGTC 427
Db 2 GATTATATTTACTACTCAAAATCAAAATCCCTAGACGACCAACGATGTC 61
QY 428 TTCAATATGTAAACAGAACAAAGTTTGTAGTACCTTAATAAGACACTCCATGAAAG 487
Db 62 TTCAATATGTAAACAGAACAAAGTTTGTAGTACCTTAATAAGACACTCCATGAAAG 121
QY 488 CTGTGATGACGTGCGCATTTGCTCTCCCATTCATTAATGCGATTACACCGGCGA 547
Db 122 CTGTGATGACGTGCGCATTTGCTCTCCCATTCATTAATGCGATTACACCGGCGA 181
QY 548 CGACAAAGCTCCGACGTAGCCCAACACCTCTAACAATGCTGCTCCGACCAAAATGG 607
Db 182 CGACAAAGCTCCGACGTAGCCCAACACCTCTAACAATGCTGCTCCGACCAAAATGG 241
QY 608 CCGACCGCTCTCTCCGACCTTCAATTTACCTCGATTCCTCTCTCTCTGCGCA 667
Db 242 CCGACCGCTCTCTCCGACCTTCAATTTACCTCGATTCCTCTCTCTCTCTGCGCA 301
QY 668 CCGACCGCTCTCTCCGACCTTCAATTTACCTCGATTCCTCTCTCTCTCTGCGCA 727
Db 302 CCGACCGCTCTCTCCGACCTTCAATTTACCTCGATTCCTCTCTCTCTCTGCGCA 361
QY 728 AGTCTCCCATCCCATTTGATTTTCAACAGTATTAAGAGCTCAACACATTTCTTAACCG 787
Db 362 AGTCTCCCATCCCATTTGATTTTCAACAGTATTAAGAGCTCAACACATTTCTTAACCG 421
QY 788 ATGAATCAGAAAGCANTGAGCTAGGCTTGTGAAACCGCGCAATTCGTTACAGG 847
Db 422 ATGAATCAGAAAGCANTGAGCTAGGCTTGTGAAACCGCGCAATTCGTTACAGG 481
QY 848 ACGACGCTTATGACGCGGAGACAGATTTCTTCAAGCTGTTGCGAACTCTGTCTAATC 907
Db 482 ACGACGCTTATGACGCGGAGACAGATTTCTTCAAGCTGTTGCGAACTCTGTCTAATC 541
QY 908 CTGCTCTAGAAAGAGTACATGAAAGTCTTCTTGAATGAAAGAGCTACAGTATCA 967
Db 542 CTGCTCTAGAAAGAGTACATGAAAGTCTTCTTGAATGAAAGAGCTACAGTATCA 601
QY 968 CTGATGTTCTTGGGATTAAGTAAATTCGATTTGGAATTAATTAAGTTCTTCGTTTAA 1027
Db 602 CTGATGTTCTTGGGATTAAGTAAATTCGATTTGGAATTAATTAAGTTCTTCGTTTAA 618
QY 1028 TTTGATGATTTGATTAAGAAAGAAAGAACTTTATCTAGTAAAGTTCTGCGGCTCTCTGT 1087
Db 619 TTTGATGATTTGATTAAGAAAGAAAGAACTTTATCTAGTAAAGTTCTGCGGCTCTCTGT 639
QY 1088 GTATTGCAAGAGGTGTGAGACTGAGATAGTTCTTCTGCGGTTGTGAGAGCTCTGTTAAG 1147
Db 640 GTATTGCAAGAGGTGTGAGACTGAGATAGTTCTTCTGCGGTTGTGAGAGCTCTGTTAAG 699
QY 1148 GAGAGGTGCTTAAGTGTGTTAAGCAAGATGAGTTTATGATGAGGCTTGCCTTCTC 1207
Db 700 GAGAGGTGCTTAAGTGTGTTAAGCAAGATGAGTTTATGATGAGGCTTGCCTTCTC 759

QY	1208	GATGCTTCGAGGAGTATGATGGACATGGAATGATCAACCGATTTTATTTACTGCTATGAGTT	1267
Db	760	GATGCTTCGAGGAGTATGATGGACATGGAATGATCAACCGATTTTATTTACTGCTATGAGTT	819
QY	1268	GTTGAGGAAGCTTTGAAAGCTTTTACAGGTAAGTTGACTTGGCTTGGTAATTTGACGACG	1327
Db	820	GTTGAGGAAGCTTTGAAAGCTTTTACAGGTAAGTTGACTTGGCTTGGTAATTTGACGACG	844
QY	1328	TTGGCTTTATTAAGAACTTTCTTGATTTGATACCTTTGTTATTTAGTCTTGTGTAGAGGAA	1387
Db	845	-----AGAGGAAA	852
QY	1388	GGAGCAATAGGCTTGCAACCGAATTTAGGTGACAAATTTGATGAGACTTTGGAAGAGATC	1447
Db	853	GGAGCAATAGGCTTGCAACCGAATTTAGGTGACAAATTTGATGAGACTTTGGAAGAGATC	912
QY	1448	ACTCCGCGCTTAATGCTTGGAGCTACTTGGCTTACCGCTTGGTGATGATTAACGTCCGAAA	1507
Db	913	ACTCCGCGCTTAATGCTTGGAGCTACTTGGCTTACCGCTTGGTGATGATTAACGTCCGAAA	972
QY	1508	AGACTAAATGGTTTAAGCGGTGTGCGGAATATTTTGTGCTGTTTGAGAGAGGTGAGCA	1567
Db	973	AGACTAAATGGTTTAAGCGGTGTGCGGAATATTTTGTGCTGTTTGAGAGAGGTGAGCA	1032
QY	1568	TCAGCTCTTGTGGGGGTTTGAACCCGTGAGAAGTTATGATGATGAGCGCTTTTACGANTG	1627
Db	1033	TCAGCTCTTGTGGGGGTTTGAACCCGTGAGAAGTTATGATGATGAGCGCTTTTACGANTG	1092
QY	1628	ACAGCTGTGACGACGAGTATACAGTTAGATACCTTTTAAATTTCTTACATATGATATA	1687
Db	1093	ACAGCTGTGACGACGAGTATACAGTTAGATACCTTTTAAATTTCTTACATATGATATA	1105
QY	1688	ACTTTAGGTTTTCTCAATTTAATGATATGTTGTGTGAGGTGATCTTTTGTAGCTACCC	1747
Db	1106	-----AGTTGATCTTTTGTAGCTACCC	1129
QY	1748	CAAGCAATATTCACACAGATCATTTGAAATTTAGAAAGTTGCACTTCTCTGTGGCTC	1807
Db	1130	CAAGCAATATTCACACAGATCATTTGAAATTTAGAAAGTTGCACTTCTCTGTGGCTC	1189
QY	1808	AAGCTTTATTTGGTATAGAACCCACACCTTTTACAGATGCTGATATGAGAAATTCACGACAC	1867
Db	1130	AAGCTTTATTTGGTATAGAACCCACACCTTTTACAGATGCTGATATGAGAAATTCACGACAC	1249
QY	1868	TTTACACAGGCTTAAGGTAATGGCTATGAGATTCCTGCGATGTTGTATGATACAAGGAATA	1927
Db	1250	TTTACACAGGCTTAAGGTAATGGCTATGAGATTCCTGCGATGTTGTATGATACAAGGAATA	1309
QY	1928	ATTGGGAATATGACTTCGGTCTAGAAAGGGAATCTGTGCACCTGTTATAGCAAAAGTTG	1987
Db	1310	ATTGGGAATATGACTTCGGTCTAGAAAGGGAATCTGTGCACCTGTTATAGCAAAAGTTG	1369
QY	1988	ATGAATGCCGATATGATGTGGGCTTAGACAGTGAAGATTCACAAATATGGAATCAGACTA	2047
Db	1370	ATGAATGCCGATATGATGTGGGCTTAGACAGTGAAGATTCACAAATATGGAATCAGACTA	1429
QY	2048	TTGTGGAATTTGTTTGGAGATTCAAATTCGTATGACAAATGATGATCTCCCTGACATAT	2107
Db	1430	TTGTGGAATTTGTTTGGAGATTCAAATTCGTATGACAAATGATGATCTCCCTGACATAT	1489
QY	2108	GCAAAATTTGTTGAAAACCTGTGTGGCAGGGGTGTCTTTCTTACAGTTACAGACACCAAG	2167
Db	1490	GCAAAATTTGTTGAAAACCTGTGTGGCAGGGGTGTCTTTCTTACAGTTACAGACACCAAG	1549
QY	2168	ATTAATAAATTTTAACTCCGGGACCTATATGATGATCTATGGTTTTTGAGTTTACTTGGAAA	2227
Db	1550	ATTAATAAATTTTAACTCCGGGACCTATATGATGATCTATGGTTTTTGAGTTTACTTGGAAA	1609
QY	2228	GAGTGGAGGTATGATCAGGGTTCTCCTTTAGCTGCTGCAACTATATGCAAGATTTGGAG	2287
Db	1610	GAGTGGAGGTATGATCAGGGTTCTCCTTTAGCTGCTGCAACTATATGCAAGATTTGGAG	1665
QY	2288	CCGAGCATATGTAAGCTATGCTATGACAGGCACTGCAGAAAGTTTTTCTTCCGCTATA	2347

Db	1610	CCGAGCATGTGAAAGCTAAGTCTATGACGGCACTGCAGAAAGTTTCTTCCCGCATATA	1729
Qy	2348	CAGATTAAGAACTCGGCTGAACCCAGAGATGCGAAGACAGTGTTAATGTAGATCTTG	2407
Db	1730	CAGATTAAGAACTCGGCTGAACCCAGAGATGCGAAGACAGTGTTAATGTAGATCTTG	1789
Qy	2408	TTTGGTAAACAATGATAGGCGGTGATAGCTGATGCTTATATGCGAAGCTGTAGAC	2467
Db	1790	TTGGTAAACAATGATAGGCGGTGATAGCTGATGCTTATATGCGAAGCTGTAGAC	1849
Qy	2468	CCTCTGAAAACTTTGAAACTAATGATTAATGCAATTCGAGCTGGGGTCTCAGAGTAGCG	2527
Db	1850	CCTCTGAAAACTTTGAAACTAATGATTAATGCAATTCGAGCTGGGGTCTCAGAGTAGCG	1909
Qy	2528	TTGATGAAACTACTGTTGAAATGTCGGTTGCTGATATGTTAAAGAGGCCAAGTGAAGA	2587
Db	1910	TTGATGAAACTACTGTTGAAATGTCGGTTGCTGATATGTTAAAGAGGCCAAGTGAAGA	1969
Qy	2588	TTCTTAGCTGCTGCTGTGTGGCAATTGAGACTGATTTCACTGTTCAAGCCAAAGATTTTCTTA	2647
Db	1970	TTCTTAGCTGCTGCTGTGTGGCAATTGAGACTGATTTCACTGTTCAAGCCAAAGATTTTCTTA	2029
Qy	2648	AAGCAGCTCATCTTTCAACGCAAGATATAGTTTTCTTCTATGGAATCTGATGCTCTA	2707
Db	2030	AAGCAGCTCATCTTTCAACGCAAGATATAGTTTTCTTCTATGGAATCTGATGCTCTA	2089
Qy	2708	CCATAGGTATGATTTAAATGATGCAATTTTCTATATCTGCAATGCTCAAAAATATGCTGT	2767
Db	2090	CCA-----CCA-----	2092
Qy	2768	TTTGTGAGCTTAAGAACATAGTCCCACTTAATACATGTCCTCAAAAATTTGACCAAGATTA	2827
Db	2093	-----	2092
Qy	2828	ACAAGTTGCTGAGTAATATTTCACTAATTAATGCTGCTGAATTTTGTGATCAAACTGTAGA	2887
Db	2093	-----	2092
Qy	2888	CAGAAATGTAATTTCACTCTCAATTTCTGTTAGAAATACGTAGGATTAAGAAATTGC	2947
Db	2093	-----	2092
Qy	2948	CTTAGTGTGCTTTGTCCAACTTTTCTTTCTTGATTTTTTCTTTTCGATTAAGGCTCA	3007
Db	2093	-----TAGGGTCA-----	2100
Qy	3008	GTCAGAGCTGACGATTCAGAAAGCACTTCCAGATGAGTGTAGACGTGCAGAAATATA	3067
Db	2101	GTCAGAGCTGACGATTCAGAAAGCACTTCCAGATGAGTGTAGACGTGCAGAAATATA	2160
Qy	3068	GTATCCAAAGTGGCAGAAAGATTAAGTCTCTGGCTTTTGGGCTGTATCAACCGCATGAAGATG	3127
Db	2161	GTATCCAAAGTGGCAGAAAGATTAAGTCTCTGGCTTTTGGGCTGTATCAACCGCATGAAGATG	2220
Qy	3128	TTACCAAGAGTGAGGGAATTAATCTACAAATCAATGTGTGAAAACTGTTGACAT	3187
Db	2221	TTTACCAAG-----	2227
Qy	3188	GATTATATAGTCTGTGCTTGTGTTGATTTCTGTTATTAAGTTTTTGGATGGCGCAATGCT	3247
Db	2228	-----AGGTTTTTGGATGGCGCAATGCT-----	2249
Qy	3248	GAAGATTTGAGTGAACAGAGCAGCTGAAACTGCGACGCTGGGTTGGTTTATGATTATAC	3307
Db	2250	GAAGATTTGAGTGAACAGAGCAGCTGAAACTGCGACGCTGGGTTTATGATTATAC	2309
Qy	3308	ACTGTTGAAACTATCTGTTGAACAAGTGTGACAGTCTCAGACAGATGAACCCGTGCTTGCT	3367
Db	2310	ACTGTTGAAACTATCTGTTGAACAAGTGTGACAGTCTCAGACAGATGAACCCGTGCTTGCT	2368
Qy	3368	GGAAGCAACTCTGAGAGAGTCTGCTGTGTTCTATCTGATTTGGTTTCAATCCAGAAACAAATGC	3427

Db 2370 GGAAGCAACTCTGAGAGTCTGCTTATCTGATTTGGTTATCCAGAAAAGATGC 2429
 Oy 3428 TACTGATGTGAGAACTTACACAAAGATGCAAGTTTCTGTCCAGTCAAGTGGAA 3487
 Db 2430 TACTGATGTGAGAACTTACACAAAGATGCAAGTTTCTGTCCAGTCAAGTGGAA 2489
 Oy 3488 AATCACTGAAGCTCTGTTCTTGTGATCATTAATTAATCTGATGATGATGATGATG 3547
 Db 2490 AATCACTGAAGCTCTGTTCTTGTGATCATTAATTAATCTGATGATGATGATGATG 2549
 Oy 3548 CGAGATCTCTTGTCTGTAATCTCTCTAGATGATGATGATGATGATGATGATGATG 3607
 Db 2550 CGAGATCTCTTGTCTGTAATCTCTCTAGATGATGATGATGATGATGATGATGATG 2609
 Oy 3608 AAAATTAACGTTCTTGTGACACCTCTTGTGATGATGATGATGATGATGATGATGATG 3667
 Db 2610 AAAATTAACGTTCTTGTGACACCAAAAAATCTTGTGATGATGATGATGATGATGATG 2669

RESULT 4
 US-10-600-070-130
 ; Sequence 130, Application US/10600070
 ; Publication No. US20040139500A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Osteoryoung, Katherine W.
 ; APPLICANT: Vitna, Stanislaw
 ; APPLICANT: Koksharova, Olga A.
 ; APPLICANT: Gao, Hongbo
 ; TITLE OF INVENTION: Plasmid Division and Related Genes and Proteins, and Methods of
 ; TITLE OF INVENTION: Use
 ; FILE REFERENCE: MSU-08153
 ; CURRENT APPLICATION NUMBER: US/10/600,070
 ; NUMBER OF SEQ ID NOS: 206
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO: 130
 ; LENGTH: 2637
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 US-10-600-070-130

Query Match 52.7%; Score 1931; DB 7; Length 2637;
 Best Local Similarity 80.4%; Pred. No. 0;
 Matches 2616; Conservative 0; Mismatches 5; Indels 632; Gaps 5;
 Oy 368 GATTAACTTATCTACTCAAAATCAAAATTCATTAACCTTGAAGCAACCAAGTCTC 427
 Db 1 GATTAACTTATCTACTCAAAATCAAAATTCATTAACCTTGAAGCAACCAAGTCTC 60
 Oy 428 TTCAATATGTAAACAGAAAGTTTGTAGTAGCTTAAAGAAAGCACTCCATGGAAG 487
 Db 61 TTCAATATGTAAACAGAAAGTTTGTAGTAGCTTAAAGAAAGCACTCCATGGAAG 120
 Oy 488 CTCTAGTCAAGTGGGATGTGTCTCTCCCACTTCAATTAGCCGATTAACACCGGGA 547
 Db 121 CTCTAGTCAAGTGGGATGTGTCTCTCCCACTTCAATTAGCCGATTAACACCGGGA 180
 Oy 548 CGACAAAGCTCGAGTAGGCAACAACCTTACAACTATCGTCCGCAAGCAATGGG 607
 Db 181 CGACAAAGCTCGAGTAGGCAACAACCTTACAACTATCGTCCGCAAGCAATGGG 240
 Oy 608 CGACCGCTCTCTCTCGCACTTCAATTTCACTCCGATTCCTCTCTCTCTCTCGCA 667
 Db 241 CGACCGCTCTCTCTCGCACTTCAATTTCACTCCGATTCCTCTCTCTCTCTCGCA 300
 Oy 668 CGGCAACCAACCGGCACTCTGTCTCTCTGCAACATTAATGATGTCGGAAGCC 727
 Db 301 CGGCAACCAACCGGCACTCTGTCTCTCTGCAACATTAATGATGTCGGAAGCC 360
 Oy 728 AGCTCCCATCCCATTAATTTCTACAGATTAATGAGTCAACCAATTTCTTAACCG 787
 Db 361 AGCTCCCATCCCATTAATTTCTACAGATTAATGAGTCAACCAATTTCTTAACCG 420

Oy 788 ATGGAATGAGAAAGCAATTCGAGTAGAGTTTCGAAAACCGCGCAATTCGGTTTCAGCG 847
 Db 421 ATGGAATGAGAAAGCAATTCGAGTAGAGTTTCGAAAACCGCGCAATTCGGTTTCAGCG 480
 Oy 848 ACGAGCTTTAATGACCGGAGAGAGATTTCTTCAAGCTGCTGGCAAACTCTGTCTAATC 907
 Db 481 ACGAGCTTTAATGACCGGAGAGAGATTTCTTCAAGCTGCTGGCAAACTCTGTCTAATC 540
 Oy 908 CTGGCTAGAGAGAGATGACATGAAGGCTCTTCTGATGATGATGATGATGATGATGATG 967
 Db 541 CTGGCTAGAGAGAGATGACATGAAGGCTCTTCTGATGATGATGATGATGATGATGATG 600
 Oy 968 CTGATGTTCTTGGAGTAAGTAATTTTCGAAATTAATTAAGTTTCTTCTGTTTAA 1027
 Db 601 CTGATGTTCTTGGAGT-----AGGTTCTGTGTCTCTGT 617
 Oy 1028 TTTCAATGATGATTAAGAGAAACCTTTATCTAGTAAGTCTCTGGGCTCTCTGT 1087
 Db 618 -----AGGTTCTGTGTCTCTGT 638
 Oy 1088 GTATTGCAAGAGGTGTGAGACTGATAGTCTTCTGGTTGTGAGGCTCTGCTTAAG 1147
 Db 639 GTATTGCAAGAGGTGTGAGACTGATAGTCTTCTGGGTTGTGAGGCTCTGCTTAAG 698
 Oy 1148 GAGAGTTGCTTAAGTCTTTAAGCAAGATGTGTTTATGATGAGCTTGTCTC 1207
 Db 699 GAGAGTTGCTTAAGTCTTTAAGCAAGATGTGTTTATGATGAGCTTGTCTC 758
 Oy 1208 GATGCTTGAGAGATGCTATGAGATTTGATGATGATGATGATGATGATGATGATGATG 1267
 Db 759 GATGCTTGAGAGATGCTATGAGATTTGATGATGATGATGATGATGATGATGATGATG 818
 Oy 1268 GTTGAGAGGCTTTGAAGCTTTTACAGGTAGTTGATCTGTTGTGTAATTTAGCAGACG 1327
 Db 819 GTTGAGAGGCTTTGAAGCTTTTAC----- 843
 Oy 1328 TTGGCTTTAAGAACTTTTGAATTTGATGATCTTTGATGATGATGATGATGATGATG 1387
 Db 844 -----AGAGGAA 851
 Oy 1388 GGAGCAAGTACCTTGCACCGGATTTACGTCAGAAATTTGATGATGATGATGATGATGATG 1447
 Db 852 GGAGCAAGTACCTTGCACCGGATTTACGTCAGAAATTTGATGATGATGATGATGATGATG 911
 Oy 1448 ACTCGCGTATGCTTGTGAGCTACTTGGCTTACCGCTTGTGATGATGATGATGATGATG 1507
 Db 912 ACTCGCGTATGCTTGTGAGCTACTTGGCTTACCGCTTGTGATGATGATGATGATGATG 971
 Oy 1508 AGACTAAATGCTTTAAGCGGTGTGCGGAATTTTGTGTGTTGAGAGAGTGAACA 1567
 Db 972 AGACTAAATGCTTTAAGCGGTGTGCGGAATTTTGTGTGTTGAGAGAGTGAACA 1031
 Oy 1568 TTAGCTCTTGTGGGGGTTTGAACCGTGAAGATTTATGATGATGATGATGATGATGATG 1627
 Db 1032 TTAGCTCTTGTGGGGGTTTGAACCGTGAAGATTTATGATGATGATGATGATGATGATG 1091
 Oy 1628 ACAGCTGCTGAGCATATACATTTAGATCTTTTAAATTTCTTACATGATATA 1687
 Db 1092 ACAGCTGCTGAGCATATACATTTAGATCTTTTAAATTTCTTACATGATATA 1104
 Oy 1688 ACTTAAGTTTCTATTTAATGATGATGATGATGATGATGATGATGATGATGATGATG 1747
 Db 1105 -----AGTTGATCTTTTGTAGTACCC 1128
 Oy 1748 CAGCAATATTCAGCAGATGATTTGATGATGATGATGATGATGATGATGATGATGATG 1807
 Db 1129 CAGCAATATTCAGCAGATGATTTGATGATGATGATGATGATGATGATGATGATGATG 1168
 Oy 1808 AAGCTTTATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1867
 Db 1189 AAGCTTTATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1248
 Oy 1868 TTGACGAGGCTAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1927

|||||
Db 1249 TTCAAGACGCTAAGCTAATGCTATGAGAGATTCCGCACTGTTGATCATACAGGATA 1308
1928 ATTGGAGATAGACTTCGGCTCTAGAAAGGGGACTCTGCGACTGCTTATAGCAAAAGTTG 1987
1309 ATTGGAGATAGACTTCGGCTCTAGAAAGGGGACTCTGCGACTGCTTATAGCAAAAGTTG 1368
Qy 1988 ATGAATGCCGATATGCTGTTGGGCTTGAACAGTGAAGATTCAAAATATAGAAATCAGCTA 2047
Db 1369 ATGAATGCCGATATGCTGTTGGGCTTGAACAGTGAAGATTCAAAATATAGAAATCAGCTA 1428
Qy 2048 TTGTGAGATTTGTTTGGAGAAATTCAAATCGTGAAGAAATGATATCCCTGACATAT 2107
Db 1429 TTGTGAGATTTGTTTGGAGAAATTCAAATCGTGAAGAAATGATATCCCTGACATAT 1488
Qy 2108 GCAAAATGTTGGAAACCTGCTGGGAGGGGTTGCTTCTCTAGTTGAGTTGAGAGAACCAAG 2167
Db 1489 GCAAAATGTTGGAAACCTGCTGGGAGGGGTTGCTTCTCTAGTTGAGAGAACCAAG 1548
Qy 2168 ATAAAAAATTTAAACTCGGGGACTATATGATGATCTATGATGATTTGAGTTACTTGAAA 2227
Db 1549 ATAAAAAATTTAAACTCGGGGACTATATGATGATCTATGATGATTTGAGTTACTTGAAA 1608
Qy 2228 GAGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2287
Db 1609 GAGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1668
Qy 2288 CCGAGCATGTGAAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2347
Db 1669 CCGAGCATGTGAAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1728
Qy 2348 CAGATGAAACTCGGCTGAAACCCAGAGATGTCAGAGAGACAGTGTGATGATGATGATGATG 2407
Db 1729 CAGATGAAACTCGGCTGAAACCCAGAGATGTCAGAGAGACAGTGTGATGATGATGATGATG 1788
Qy 2408 TTGTGAAACAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2467
Db 1789 TTGTGAAACAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1848
Qy 2468 CCTCGAAACCTTTGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2527
Db 1849 CCTCGAAACCTTTGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1908
Qy 2528 TTGATGAAACTACTGTTGAAATGTCCTGCTGATGATGATGATGATGATGATGATGATGATG 2587
Db 1909 TTGATGAAACTACTGTTGAAATGTCCTGCTGATGATGATGATGATGATGATGATGATGATG 1968
Qy 2588 TCCCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2647
Db 1969 TCCCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2028
Qy 2648 AAAGCAGCTCACTTTTCAACGCAAGATGATGATGATGATGATGATGATGATGATGATGATG 2707
Db 2029 AAAGCAGCTCACTTTTCAACGCAAGATGATGATGATGATGATGATGATGATGATGATGATG 2088
Qy 2708 CCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2767
Db 2089 CCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2091
Qy 2768 TTTGTGAGCTAGAGACATAGTTCCCACTTAATACATGTCACAAAAGTTGTACCAAGATTA 2827
Db 2092 TTTGTGAGCTAGAGACATAGTTCCCACTTAATACATGTCACAAAAGTTGTACCAAGATTA 2091
Qy 2828 ACAAGTGTCTGAGTAATTTTCACTAATATATGCTGCTGATGATTTTGTATCAAACTGTAGA 2887
Db 2092 ACAAGTGTCTGAGTAATTTTCACTAATATATGCTGCTGATGATTTTGTATCAAACTGTAGA 2091
Qy 2888 CAGAAATGTAATTTTCACTCTCAACATTTCTGTTAGATAACGTAGAGATTAGAGATTGC 2947
Db 2092 CAGAAATGTAATTTTCACTCTCAACATTTCTGTTAGATAACGTAGAGATTAGAGATTGC 2091
Qy 2948 CTTAGTGTGGCTTTGCCAATTTTCTTCTTGAATTTTCTTGTGATGATGATGATGATGATGATG 3007
Db 2092 CTTAGTGTGGCTTTGCCAATTTTCTTCTTGAATTTTCTTGTGATGATGATGATGATGATGATG 2091

Db 2092 -----TAGGCTCA 2099
Qy 3008 GTCAAGCTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3067
Db 2100 GTCAAGCTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2159
Qy 3068 GTATCAAGTGTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3127
Db 2160 GTATCAAGTGTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2219
Qy 3128 TTACCAAGTGTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3187
Db 2220 TTACCAAGTGTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2226
Qy 3188 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3247
Db 2227 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2248
Qy 3248 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3307
Db 2249 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2308
Qy 3308 ACTGTTGAACTATCTGTTGACAGTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTGAC 3367
Db 2309 ACTGTTGAACTATCTGTTGACAGTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTGAC 2368
Qy 3368 GAAAGCACTCTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3427
Db 2369 GAAAGCACTCTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2428
Qy 3428 TACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3487
Db 2429 TACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2488
Qy 3488 AATCAGTGAAGGCTCTGTTCTGATCATATATATCATATGATGATGATGATGATGATGATGATG 3547
Db 2489 AATCAGTGAAGGCTCTGTTCTGATCATATATATCATATGATGATGATGATGATGATGATGATG 2548
Qy 3548 CGAGATTCCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 3607
Db 2549 CGAGATTCCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2608
Qy 3608 AAAATTAACGTTTC 3620
Db 2609 AAAATTAACGTTTC 2621

RESULT 5
US-10-600-070-1
; Sequence 1, Application US/10600070
; Publication No. US20040139500A1
; GENERAL INFORMATION:
; APPLICANT: Oseireyoung, Katherine W.
; APPLICANT: Vitsha, Stanislaw
; APPLICANT: Koksharova, Olga A.
; TITLE OF INVENTION: Placid Division and Related Genes and Proteins, and Methods of
; FILE REFERENCE: MSU-08153
; CURRENT FILING DATE: 2003-06-20
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 2406
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-600-070-1

Query Match 47.0%; Score 1724; DB 7; Length 2406;
Best Local Similarity 79.2%; Pred. No. 0;
Matches 2406; Conservative 0; Mismatches 0; Indels 632; Gaps 5;

481 ATGGAAGCTCTGAGTCAAGTCCGGAATTGGTCTCTCCCAATTCATATATGCCGATTACA 540
Db 1 ATGGAAGCTCTGAGTCAAGTCCGGAATTGGTCTCTCCCAATTCATATATGCCGATTACA 60
OY 541 CCGGAGACGACAAAGCTCCGAGTAGCCACAAACACTTACAACTATCTGCTCCGACG 600
Db 61 CCGGAGACGACAAAGCTCCGAGTAGCCACAAACACTTACAACTATCTGCTCCGACG 120
OY 601 AAATGGGCGGACCGTCTCTCCGACTTCGAAATTCACCTCCGATTCCTCTCTCC 660
Db 121 AAATGGGCGGACCGTCTCTCCGACTTCGAAATTCACCTCCGATTCCTCTCTCC 180
OY 661 TTCGCAACCGGACCAACCAACCGGCACTCTGCTCTCTGCACTCTATATGATGCTCC 720
Db 181 TTCGCAACCGGACCAACCAACCGGCACTCTGCTCTCTGCACTCTATATGATGCTCC 240
OY 721 GAAGCGACGCTCCCAATCCCAATGATTCTACAGGTATAGAGTCAAAACATTTTC 780
Db 241 GAAGCGACGCTCCCAATCCCAATGATTCTACAGGTATAGAGTCAAAACATTTTC 300
OY 781 TTAAACGATGGAATGAGAAAGACATTCGAGCTAGGGTTTCGAAACCGCGCAATTCGCT 840
Db 301 TTAAACGATGGAATGAGAAAGACATTCGAGCTAGGGTTTCGAAACCGCGCAATTCGCT 360
OY 841 TTCACGACGACGCTTTTAATCAGCCGAGACAGATTTCTTGAAGCTGCTTGGCAACTCTG 900
Db 361 TTCACGACGACGCTTTTAATCAGCCGAGACAGATTTCTTGAAGCTGCTTGGCAACTCTG 420
OY 901 TCTATCTCGGCTAGAAAGAGTACATGAAAGCTCTTGATGATGAAAGACTCA 960
Db 421 TCTATCTCGGCTAGAAAGAGTACATGAAAGCTCTTGATGATGAAAGACTCA 480
OY 961 GTCATCATGATGTTCTTTGGAGTAAGTAAATTCGATTCGGAATTAATAAGTTCTTC 1020
Db 481 GTCATCATGATGTTCTTTGGAGT----- 504
OY 1021 GTTTTAATTTATGAAATGGAATAAGAAAGAACTTTATCTAGTGAAAGCTCTCGGAGC 1080
Db 505 -----AGGTTCTCGGAGC 518
OY 1081 TCTCTGTGATTTGAAAGAGTGTGAGACGTGAGATGATCTCTGGGTTGGAGGCTCT 1140
Db 519 TCTCTGTGATTTGAAAGAGTGTGAGACGTGAGATGATCTCTGGGTTGGAGGCTCT 578
OY 1141 GCTTAAGAGAGGTGCTTAAGTCTTTAAGCAAGATGATGTTTATAGTTAGCCCTTGC 1200
Db 579 GCTTAAGAGAGGTGCTTAAGTCTTTAAGCAAGATGATGTTTATAGTTAGCCCTTGC 638
OY 1201 GTTCTCGATGTCTGAGGAGTGTATGCAATTCGATCCACTGATTTTATTAATCTGTTA 1260
Db 639 GTTCTCGATGTCTGAGGAGTGTATGCAATTCGATCCACTGATTTTATTAATCTGTTA 698
OY 1261 TGAGTTTGTGAGAACTTTGAACTTTTAC----- 730
Db 699 TGAGTTTGTGAGAACTTTGAACTTTTAC----- 730
OY 1321 ACGAGCGTGGCTTATTAAGAACTTTCTTGAATTTATATCTTGTATGAGCTGTGTA 1380
Db 731 -----A 731
OY 1381 GGAGGAAGAGCAAGTATGCTTGACCCGAGATTAAGTGCACAAAATGATGAGACTTTGGA 1440
Db 732 GGAGGAAGAGCAAGTATGCTTGACCCGAGATTAAGTGCACAAAATGATGAGACTTTGGA 791
OY 1441 AGAGATCACTCCGCTTATGCTTGAAGTACTTGGCTTACCGCTTGGTGTATGATTAAGC 1500
Db 792 AGAGATCACTCCGCTTATGCTTGAAGTACTTGGCTTACCGCTTGGTGTATGATTAAGC 851
OY 1501 TGCGAAAAGACTAAATGCTTTAAGCGGTGTGCGAAATATTTTGTGCTGTGTGAGAGG 1560
Db 852 TGCGAAAAGACTAAATGCTTTAAGCGGTGTGCGAAATATTTTGTGCTGTGTGAGAGG 911
OY 1561 TGAGCATCAAGCTCTTGTGGGGGTTTGAACCCGTGAGAAAGTTTATGAATGAGGCGTTT 1620

Db 912 TGAGCATCAAGCTCTTGTGGGGGTTTGAACCCGTGAGAAAGTTTATGAATGAGGCGTTT 971
OY 1621 ACGATGACAGCTGTGAGAGGTATACAGTTTATACATCTTTTATTTTATTTTATGACA 1680
Db 972 ACGATGACAGCTGTGAG----- 991
OY 1681 TGATATTAATTTAGGTTTCTATTTTAATGATATGATGTGTGAGGTTGATCTTTTGT 1740
Db 992 -----AGGTTGATCTTTTGT 1008
OY 1741 GCTACCCCAAGCAATATTCAGAGAGTCAATTTGAAGTTTGAAGTTTGAAGTTTGAAGTTT 1800
Db 1009 GCTACCCCAAGCAATATTCAGAGAGTCAATTTGAAGTTTGAAGTTTGAAGTTTGAAGTTT 1068
OY 1801 GTGGCTCAAGCTTTTATGTTAAGAGCAACCTTTTACAGAGTGTATGAGCAATTC 1860
Db 1069 GTGGCTCAAGCTTTTATGTTAAGAGCAACCTTTTACAGAGTGTATGAGCAATTC 1128
OY 1861 CAGCAACTTCAGAGGCTAAGGTATAGCTATGAGATTCCTGCGATGTTGATATACA 1920
Db 1129 CAGCAACTTCAGAGGCTAAGGTATAGCTATGAGATTCCTGCGATGTTGATATACA 1188
OY 1921 CGGAATTAATGGGAGATGACTTGGCTTAAAGAGGAGCTCTGTGCACTGCTTATAGGC 1980
Db 1189 CGGAATTAATGGGAGATGACTTGGCTTAAAGAGGAGCTCTGTGCACTGCTTATAGGC 1248
OY 1981 AAAGTTGATGAAATCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2040
Db 1249 AAAGTTGATGAAATCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1308
OY 2041 CCAAGTATTTGAGATTTGTTTGGAGAAATCAATCTGATGATGATGATGATGATGATGAT 2100
Db 1309 CCAAGTATTTGAGATTTGTTTGGAGAAATCAATCTGATGATGATGATGATGATGATGAT 1368
OY 2101 GCACTATGCAATTTGTTGAAACCTGTTGGCAAGGTTGTTCTTCTAGGTTCAAGAC 2160
Db 1369 GCACTATGCAATTTGTTGAAACCTGTTGGCAAGGTTGTTCTTCTAGGTTCAAGAC 1428
OY 2161 ACCAAAGTAAATAATTTAACTCGGGGACTACTATGATGATGATGATGATGATGATGATGAT 2220
Db 1429 ACCAAAGTAAATAATTTAACTCGGGGACTACTATGATGATGATGATGATGATGATGAT 1488
OY 2221 TTGGAAGAGTGAAGGTATCAGGGTCTCTTAACTGCTGCTGCAATATGAGCAGG 2280
Db 1489 TTGGAAGAGTGAAGGTATCAGGGTCTCTTAACTGCTGCTGCAATATGAGCAGG 1548
OY 2281 ATTGAAGCCGAGATGTGAAGTATGCTATGCAAGCACTGCAAGAAAGTTTCTCTCC 2340
Db 1549 ATTGAAGCCGAGATGTGAAGTATGCTATGCAAGCACTGCAAGAAAGTTTCTCTCC 1608
OY 2341 CGCTATACAGTAAACTCGGCTGAACCCCAAGATGTGCAAGAGACAGTGTATAGTTA 2400
Db 1609 CGCTATACAGTAAACTCGGCTGAACCCCAAGATGTGCAAGAGACAGTGTATAGTTA 1668
OY 2401 GATCTGTGTGTAACAATGTAAGCCGATGATGAGCTGTGTCTTATATGCAAGACT 2460
Db 1669 GATCTGTGTGTAACAATGTAAGCCGATGATGAGCTGTGTCTTATATGCAAGACT 1728
OY 2461 GTAAAGCCCTTGAAAACCTTTGAAACTATATGATTTATGCAATTCGAGCTGGGCTCAGAG 2520
Db 1729 GTAAAGCCCTTGAAAACCTTTGAAACTATATGATTTATGCAATTCGAGCTGGGCTCAGAG 1788
OY 2521 AGTAGCGTGAATGAACCTAGCTTGAATGTCGTCGATATAGTTAAAGAGGCAAGT 2580
Db 1789 AGTAGCGTGAATGAACCTAGCTTGAATGTCGTCGATATAGTTAAAGAGGCAAGT 1848
OY 2581 GTGAAGATCTAGCTGTGTGTGTGCAATGGAATGGAATTTCACTGTTCAAGCAGAAATAT 2640
Db 1849 GTGAAGATCTAGCTGTGTGTGTGCAATGGAATGGAATTTCACTGTTCAAGCAGAAATAT 1908
OY 2641 TTTCTTAAAGACAGCTCATCTTTTCAACGCAAGATATGTTTCTTATGGAATCTGAT 2700

```
Db 1909 TTTCTTAAAGACGCTCATCTTTTCAACGACGAGTATGTTTCTTCTATGGAATCTGAT 1968
Qy 2701 GTGCTACCATAGATGATGATTAAATGATGCAATTTTCATATATCTGATCTCTCAAAATA 2760
Db 1969 GTGCTACCA----- 1978
Qy 2761 TGCTTTGTTTGTAGCTAGAACATATGTTCCCATTAATACATGTCCTCAAAAGTTTACC 2820
Db 1979 ----- 1978
Qy 2821 AAGATTACAAGTTGCTGATGTAATTTTCACTAATATGCTGCTGAATTTTGTATCAAA 2880
Db 1979 ----- 1978
Qy 2881 CTGTAGACAGAAATGTAATTTTCACTCTCAATTTCTGTTTGAATAACGTAGATTAG 2940
Db 1979 ----- 1978
Qy 2941 AGATTGCTTATGTTGCTGCTTGTCTTCCACTTTTCTTCTGATTTTCTTTCGATTT 3000
Db 1979 ----- 1979
Qy 3001 AGGCTCACTGACAGCTGACGATTCGAGACCACTTCCGAAATGATGCTAGACTGACGA 3060
Db 1980 AGGCTCACTGACAGCTGACGATTCGAGACCACTTCCGAAATGATGCTAGACTGACGA 2039
Qy 3061 GAATATAGTATTCAGATGAGCAGAAAGATTAACTCTGCTTTTGGGCTCTGACACCGAT 3120
Db 2040 GAATATAGTATTCAGATGAGCAGAAAGATTAACTCTGCTTTTGGGCTCTGATCACCGAT 2099
Qy 3121 AGAAATGTTACCAAGGTGAGGGAATTAATCTAATCAATCAATGTTGTGAAAACCTGT 3180
Db 2100 AGAAATGTTACCAAG----- 2113
Qy 3181 TGGAATATTAATGCTGATGCTGCTGCTTGTGATTCGTTATTAAGTTTGTGATGGGC 3240
Db 2114 -----AGGTTTGTGATGGGC 2128
Qy 3241 GAATGCTGAGAGTTTGAATGACGACGAGCTGAAACTGCGAGCTTGGTGGTTATG 3300
Db 2129 GAATGCTGAGAGTTTGAATGACGACGAGCTGAAACTGCGAGCTTGGTGGTTATG 2188
Qy 3301 ATTATACCTGTTGAAACTATCTGTTGACAGTGTACAGTCTGACAGATGGAACCCGTG 3360
Db 2189 ATTATACCTGTTGAAACTATCTGTTGACAGTGTACAGTCTGACAGATGGAACCCGTG 2248
Qy 3361 CTCTGCTGAGAGCACTCTGAGAGAGTCTGCTGCTATCTGATTTGGTTATCCAGAA 3420
Db 2249 CTCTGCTGAGAGCACTCTGAGAGAGTCTGCTGCTATCTGATTTGGTTATCCAGAA 2308
Qy 3421 ACAATGCTACTGATGTCAGAACTCTACACAAGATGAAAGTTTCTGTCAGATGACG 3480
Db 2309 ACAATGCTACTGATGTCAGAACTCTACACAAGATGAAAGTTTCTGTCAGATGACG 2368
Qy 3481 GGTGAAAATCATCTGAGAGCTCTGTTCTTGATCATATAA 3518
Db 2369 GGTGAAAATCATCTGAGAGCTCTGTTCTTGATCATATAA 2406
```

```
RESULT 6
US-10-600-070-9
; Sequence 9, Application US/10600070
; Publication No. US20040139500A1
; GENERAL INFORMATION:
; APPLICANT: Osteoryoung, Katherine W.
; APPLICANT: Vitna, Stanislaw
; APPLICANT: Kokeharova, Olga A.
; APPLICANT: Gao, Hong
; TITLE OF INVENTION: Plasmid Division and Related Genes and Proteins, and Methods of
; FILE REFERENCE: MSU-08153
; CURRENT APPLICATION NUMBER: US/10/600,070
; CURRENT FILING DATE: 2003-06-20
```

```
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 9
; LENGTH: 2406
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-600-070-9
```

```
Query Match 46.9%; Score 1720.8; DB 7; Length 2406;
Best Local Similarity 79.1%; Pred. No. 0;
Matches 2404; Conservative 0; Mismatches 2; Indels 632; Gaps 5;
```

```
Qy 481 ATGGAAGCTCTGATCAGCTGCGATGCTGCTCTCCCATTCCAATATGCCATTTACA 540
Db 1 ATGGAAGCTCTGATCAGCTGCGATGCTGCTCTCCCATTCCAATATGCCATTTACA 60
Qy 541 CCGGCGACGACAAAGCTCCGACGAGCACAACCTCTCAACTATCTGCTCGCAGC 600
Db 61 CCGGCGACGACAAAGCTCCGACGAGCACAACCTCTCAACTATCTGCTCGCAGC 120
Qy 601 AATGGGCGACGCTCTTCTCTCCGACTTCAATTTCACTCCGATTCCTCTCTCC 660
Db 121 AATGGGCGACGCTCTTCTCTCCGACTTCAATTTCACTCCGATTCCTCTCTCC 180
Qy 661 TTGCGCACCGCACCAACCAACCGCACTCTGCTCTCTGCAACATTTATGATGCTCC 720
Db 181 TTGCGCACCGCACCAACCAACCGCACTCTGCTCTCTGCAACATTTATGATGCTCC 240
Qy 721 GAAGCGACGCTCCCATCCCATTTGATTTGACAGGATTTAGAGCTCAAAACATTTG 780
Db 241 GAAGCGACGCTCCCATCCCATTTGATTTGACAGGATTTAGAGCTCAAAACATTTG 300
Qy 781 TTAAACGATGAAATCAGAAAGACATTCGAAAGCTTTCGAAACCGCGCAATTCGGT 840
Db 301 TTAAACGATGAAATCAGAAAGACATTCGAAAGCTTTCGAAACCGCGCAATTCGGT 360
Qy 841 TTACGCGACGACGCTTAAATCAACCGGACAGATTTCTTCAAGCTCTTCCGAAACTCTG 900
Db 361 TTACGCGACGACGCTTAAATCAACCGGACAGATTTCTTCAAGCTCTTCCGAAACTCTG 420
Qy 901 TCTAATCTCGGCTTAGAAGAGTACATGAAGGCTCTCTGATGATGAAGAGCTACA 960
Db 421 TCTAATCTCGGCTTAGAAGAGTACATGAAGGCTCTCTGATGATGAAGAGCTACA 480
Qy 961 GTCATCACTGATGTTCTTGGGATTAAGTAAATTCGATTCGGAATTAAGTTTCTTC 1020
Db 481 GTCATCACTGATGTTCTTGGGAT----- 504
Qy 1021 GTTTTAATTCATGAAATGGAATTAAGAGAACTTTTATCTAGTAAAGTTCTGGGCG 1080
Db 505 -----AAGTTCTCTGGGCG 518
Qy 1081 TCTCTGATATTCGAAGAAGGTGTGAGACTGATGTTCTTCGGGTGTGAGGCTCT 1140
Db 519 TCTCTGATATTCGAAGAAGGTGTGAGACTGATGTTCTTCGGGTGTGAGGCTCT 578
Qy 1141 GCTTAAGAGAGGTGTGCTTAAGTCTTTAAGCAAGATGTGTTTAACTTAAGCGCTTGC 1200
Db 579 GCTTAAGAGAGGTGTGCTTAAGTCTTTAAGCAAGATGTGTTTAACTTAAGCGCTTGC 638
Qy 1201 GTTTCGTGATGTCGAGGAGTGTATGATGATGATGATGATGATGATGATGATGATG 1260
Db 639 GTTTCGTGATGTCGAGGAGTGTATGATGATGATGATGATGATGATGATGATGATG 698
Qy 1261 TGAATTTTGAAGAACTTTGAAGCTTTTAAAGCTTTTAAAGCTTTTAAAGCTTTTAA 1320
Db 699 TGAATTTTGAAGAACTTTGAAGCTTTTAAAGCTTTTAAAGCTTTTAAAGCTTTTAA 730
Qy 1321 ACGAGCGTGGCTTTAATAAGAACTTTTGAATTTGATGATGATGATGATGATGATG 1380
Db 731 -----A 731
Qy 1381 GGAGGAAGAGCAAGTAGCTTGACCGGATTTAAGTGACAAATGATGAGACTTTGGA 1440
```



```

Sequence 128 Application US/10600070
Publication No. US20040139500A1
GENERAL INFORMATION:
APPLICANT: Osteryoung, Katherine W.
APPLICANT: Vitna, Stanislaw
APPLICANT: Kokcharova, Olga A.
APPLICANT: Gao, Hongbo
TITLE OF INVENTION: Plastid Division and Related Genes and Proteins, and Methods of
TITLE OF INVENTION: Use
FILE REFERENCE: MSU-08153
CURRENT APPLICATION NUMBER: US/10/600,070
CURRENT FILING DATE: 2003-06-20
NUMBER OF SEQ ID NOS: 206
SOFTWARE: PatentIn version 3.2
SEQ ID NO 128
LENGTH: 2406
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-10-600-070-128

```

Query Match	46.8%	Score 1717.6;	DB 7;	Length 2406;
Best Local Similarity	79.1%;	Pred. No. 0;		
Matches 2402;	Conservative	0;	Mismatches 4;	Indels 632;
				Gaps 5;

481 |||GGGAAGCTCTGAGTACGTCGGGCAATTGGTCTCTCCCATTCCTCAATATAGCCGATTTACCA 540

Db 1 ATGGAAGCTCTGATGTACGTCGGGCAATTGGTCTCTCCCATTCCTCAATATAGCCGATTTACCA 60

Qy 541 CCGGCGACGACAAAGCTCCGACGTAGCCACAAACGCTCTCAACTATCTGCTCCGCGAGC 600

Db 61 CCGGCGACGACAAAGCTCCGACGTAGCCACAAACGCTCTCAACTATCTGCTCCGCGAGC 120

Qy 601 AAATGGGCGGACCGCTCTTCTCTCCGACTTCGAATTTCACTCCGATTCCTCTCTCTCC 660

Db 121 AAATGGGCGGACCGCTCTTCTCTCCGACTTCGAATTTCACTCCGATTCCTCTCTCTCC 180

Qy 661 TTTCGCCACCGGCCACCAACCAACGCGACCTCTGTCTCTTCGCCACACTATATGATCGTCCC 720

Db 181 TTTCGCCACCGGCCACCAACCAACGCGACCTCTGTCTCTTCGCCACCATATATGATCGTCCC 240

Qy 721 GAAGGCGACGTCGCCCATCCCATTTGATTTCTTACCAAGGATTTAGGAGCTCAACACATTTCC 780

Db 241 GAAGGCGACGTCGCCCATCCCATTTGATTTCTTACCAAGGATTTAGGAGCTCAACACATTTCC 300

Qy 781 TTAAACGATGGAATCAGAAAGAGCATTCGAAAGTAGGGTTTTCGAAACCGCGCAATTCGGT 840

Db 301 TTAAACGATGGAATCAGAAAGAGCATTCGAAAGTAGGGTTTTCGAAACCGCGCAATTCGGT 360

Qy 841 TTCAAGCGACGACGCTTTTAATCAAGCCGAGACAGATTCTTCAAGCTGCTTGGGAAACTCTG 900

Db 361 TTCAAGCGACGACGCTTTTAATCAAGCCGAGACAGATTCTTCAAGCTGCTTGGGAAACTCTG 420

Qy 901 TCTAATCCCTCGGCTCTAGGAAGAGGTACAATGAAGAGGCTTCCTGTGATGATGAAGAAGCTACA 960

Db 421 TCTAATCCCTCGGCTCTAGGAAGAGGTACAATGAAGAGGCTTCCTGTGATGATGAAGAAGCTACA 480

Qy 961 GTCATCACTGATGTTCCTTGGGATAAGGTAAATTTCGATTTCCGATAATATAAAGTTTCTTC 1020

Db 481 GTCATCACTGATGTTCCTTGGGATAAGGTAAATTTCGATTTCCGATAATATAAAGTTTCTTC 504

Qy 1021 GTTTTAATTTCAATGAATTGATTAAGGAAGAACTTTATCTAGTGAAGGTTCTCTGCGGC 1080

Db 505 -----AAGGTTCTGTGTC 518

Qy 1081 TCTCTGTGTATTGCAAGAGGTGTGTAGACTGAGATAGTTCTTCGGGTGTGTAGGCTCT 1140

Db 519 TCTCTGTGTATTGCAAGAGGTGTGTAGACTGAGATAGTTCTTCGGGTGTGTAGGCTCT 578

Qy 1141 GCTTAAGGAGAGGTTGCTTAAGTGGTTTAAGCAAGATGTGTATTAGTTATAGGCGCTTGC 1200

Db 579 GCTTAAGGAGAGGTTGCTTAAGTGGTTTAAGCAAGATGTGTATTAGTTATAGGCGCTTGC 638

Qy 1201 GTTTCCTCATGTCTCGAGGGATGCTATGGCAATGGATCCACTGATTTTATTACGTGTTA 1260

Db	639	GTTCCTCGATGTCCTCGAGGGATGCTATGCAATTTGATCCACTGATTTTAACTGGTTA	658
Qy	1261	TGAGTTTGTGAGAGACCTTTGAAAGCTTTTACAGGTAGTTGA	1320
Db	699	TGAGTTTGTGAGAGACCTTTGAAAGCTTTTAC	730
Qy	1321	ACGAGCGTTGGCTTTATTAAGACCTTCTTGATTTGATTA	1380
Db	731	-----A 731	
Qy	1381	GGAGAAAGAGCAAGTAGTAGCTTGACCCGATTTTACGNGACAATTTGATGACA	1440
Db	732	GGAGAAAGAGCAAGTAGTAGCTTGACCCGATTTTACGNGACAATTTGATGACA	791
Qy	1441	AGAGATCACTCCGCGTTATGTCCTTGAGACTACTTGCGTTACCGCTTGGTAGATTA	1500
Db	792	AGAGATCACTCCGCGTTATGTCCTTGAGACTACTTGCGTTACCGCTTGGTAGATTA	851
Qy	1501	TGCGAAAAGACTAATAGTTTAAAGCGGTGTCGGAAATTTTGTGGTCTGTGAGAGAG	1560
Db	852	TGCGAAAAGACTAATAGTTTAAAGCGGTGTCGGAAATTTTGTGGTCTGTGAGAGAG	911
Qy	1561	TGAGCAATCAAGCTTGTGTGGGGGTTGACCCCGAGAACTTATGATGAATAGCGTTT	1620
Db	912	TGAGCAATCAAGCTTGTGTGGGGGTTGACCCCGAGAACTTATGATGAATAGCGTTT	971
Qy	1621	ACGATGACAGCTGCTGAGAGAGTATCAAGTTTAAATACCTTTTAAATTTCTTTAGCA	1680
Db	972	ACGATGACAGCTGCTGAGAGAGTATCAAGTTTAAATACCTTTTAAATTTCTTTAGCA	991
Qy	1681	TGATTAATCTTAAAGTTTCTCATATTTAATGATGTGTGAGAGTTGATCTTTTGTGA	1740
Db	992	-----AGGTTGATCTTTTGTGA	1008
Qy	1741	GCTACCCCAAGCAATATTCCAGAGAGTCAATTTGAATTTAGAAATGTCAGCTTCTT	1800
Db	1009	GCTACCCCAAGCAATATTCCAGAGAGTCAATTTGAATTTAGAAATGTCAGCTTCTT	1068
Qy	1801	GTGACTCAAGCTTTTATTTGTTAAGAAACCAACCTTTTACAGATGCTGATTAAGCAATTC	1860
Db	1069	GTGACTCAAGCTTTTATTTGTTAAGAAACCAACCTTTTACAGATGCTGATTAAGCAATTC	1128
Qy	1861	CAGCAACTTCAGCAGGCTAAGTATAGGCTATGAGAGATTCCTGCGATGTTGATATGATACA	1920
Db	1129	CAGCAACTTCAGCAGGCTAAGTATAGGCTATGAGAGATTCCTGCGATGTTGATATGATACA	1188
Qy	1921	CGGAATTAATTGGAGATAGACTTCGCGTCTAGAAAGGGGACTCTGTGCACTGCTTATAGGC	1980
Db	1189	CGGAATTAATTGGAGATAGACTTCGCGTCTAGAAAGGGGACTCTGTGCACTGCTTATAGGC	1248
Qy	1981	AAAGTTGATGATGCTCGTATGTGGTGTGGGCTTGAACAGTAGAGATTCACAATATAGGAAT	2040
Db	1249	AAAGTTGATGATGCTCGTATGTGGTGTGGGCTTGAACAGTAGAGATTCACAATATAGGAAT	1308
Qy	2041	CCAGCTATTTGTGAGTTTGTTTTGAGAAATTCAAATCGTATGACAAATGATGATCTTCCT	2100
Db	1309	CCAGCTATTTGTGAGTTTGTTTTGAGAAATTCAAATCGTATGACAAATGATGATCTTCCT	1368
Qy	2101	GGACTATGCAAAATTTGTGAAAACCTGTGTTGGCAGGGGTGTCTTCTCTAGGTTCAAGAC	2160
Db	1369	GGACTATGCAAAATTTGTGAAAACCTGTGTTGGCAGGGGTGTCTTCTCTAGGTTCAAGAC	1428
Qy	2161	ACCAAAAGTAAAAAATTTAAACTCGGGGACCTACATGATATATCTATAGTTTGAAGTTAC	2220
Db	1429	ACCAAAAGTAAAAAATTTAAACTCGGGGACCTACATGATATATCTATAGTTTGAAGTTAC	1488
Qy	2221	TTGGAAGAAGTGAAGTAGTTTCAAGGTTCTCTTTAGCTGCTGCTGCACTATGCGAAG	2280
Db	1489	TTGGAAGAAGTGAAGTAGTTTCAAGGTTCTCTTTAGCTGCTGCTGCACTATGCGAAG	1548
Qy	2281	ATTGAGCCGAGCATGTGAAGCTAATGCTATGTCAGGCACTGCAAAAATTTTCTCTTC	2340

Db 1549 ATTGAGCCGAGATGTGAAGCTAGTCTATGACAGCACTGCAAGATTTTCTTCC 1608
 Qy 2341 CGCTATACAGATGAAGAACTCGGCTGAAGCCAGAGATGTGCAGAGACAGTGTGTAGTGA 2400
 Db 1609 CGCTATACAGATGAAGAACTCGGCTGAAGCCAGAGATGTGCAGAGACAGTGTGTAGTGA 1668
 Qy 2401 GATCCGTGTGTGAACAATGAGGCGGTGATGTGAGCCGTGTGTCTTTATTGCAAGAGCT 2460
 Db 1669 GATCCGTGTGTGAACAATGAGGCGGTGATGTGAGCCGTGTGTCTTTATTGCAAGAGCT 1728
 Qy 2461 GTAAAGACCTCTGAAAACTTTGAAACTAAATGATTAATGCAATTCGAGCTGAGGCTGAGAG 2520
 Db 1729 GTAAAGACCTCTGAAAACTTTGAAACTAAATGATTAATGCAATTCGAGCTGAGGCTGAGAG 1788
 Qy 2521 AGTAGCGTTGATGAACCTAATGTTGAATGTCGCTGTGTGTGAATGTTAAAGAGGCAAGT 2580
 Db 1789 AGTAGCGTTGATGAACCTAATGTTGAATGTCGCTGTGTGTGAATGTTAAAGAGGCAAGT 1848
 Qy 2581 GTGAAGATCTGAGCTGCTGT 2640
 Db 1849 GTGAAGATCTGAGCTGCTGT 1908
 Qy 2641 TTTCTTAAAGAGAGCTCATCTTTTCAAGCAGAGATAGTGTCTTCTATGAGATCTGAT 2700
 Db 1909 TTTCTTAAAGAGAGCTCATCTTTTCAAGCAGAGATAGTGTCTTCTATGAGATCTGAT 1968
 Qy 2701 GTGCTACCATAGATGATTAATGATGATCAATTTTCAATATCTGATTCGCAAAATA 2760
 Db 1969 GTGCTACCA----- 1978
 Qy 2761 TCTTGT 2820
 Db 1979 ----- 1978
 Qy 2821 AAGATTAACAAGTTGT 2880
 Db 1979 ----- 1978
 Qy 2881 CTGTAGACAGAAATGTAATTTCACTGTCAATTTCTGTGTGTGTGTGTGTGTGTGTGTGT 2940
 Db 1979 ----- 1978
 Qy 2941 AGATTGCCCTTGT 3000
 Db 1979 ----- 1979
 Qy 3001 AGGTCAGTCAAGGCTGAGATTCAGAGCACTTCCAGAAATGATGCTAGAGATCTGCAGA 3060
 Db 1980 AGGTCAGTCAAGGCTGAGATTCAGAGCACTTCCAGAAATGATGCTAGAGATCTGCAGA 2039
 Qy 3061 GAATATGATTCAGAGGCTGAGATTTAATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3120
 Db 2040 GAATATGATTCAGAGGCTGAGATTTAATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2099
 Qy 3121 AGAAATGTTACAGAGGCTGAGATTTAATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3180
 Db 2100 AGAAATGTTACAG----- 2113
 Qy 3181 TGGAATGATTTATGCTGT 3240
 Db 2114 ----- 2128
 Qy 3241 GAATGTCAGATTTGATGCTGACAGAGCTGAAGCTGAGGCTGTGTGTGTGTGTGTGTGTGTGT 3300
 Db 2129 GAATGTCAGATTTGATGCTGACAGAGCTGAAGCTGAGGCTGTGTGTGTGTGTGTGTGTGTGT 2188
 Qy 3301 ATTATACACTGTGTGAACCTATCTGTGTGAAGTGTGACAGTCTCAGCAAGTGAAGCCGTG 3360
 Db 2189 ATTATACACTGTGTGAACCTATCTGTGTGAAGTGTGACAGTCTCAGCAAGTGAAGCCGTG 2248
 Qy 3361 CTGTGTGTGAAGCACTGTGAGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3420
 Db 2249 CTGTGTGTGAAGCACTGTGAGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2308

Qy 3421 ACAATGCTACTGATGTGCAAACTTACACAAAGATATGAGATTTTCTGTCCAAGTCAG 3480
 Db 2309 ACAATGCTACTGATGTGCAAACTTACACAAAGATATGAGATTTTCTGTCCAAGTCAG 2368
 Qy 3481 GGTGGAATCACTGAGAGGCTTGTCTTGTGATCAATA 3518
 Db 2369 GGTGGAATCACTGAGAGGCTTGTCTTGTGATCAATA 2406
 RESULT 8
 US-10-600-070-132/c
 ; Sequence 132, Application US/10600070
 ; Publication No. US20040139500a1
 ; GENERAL INFORMATION:
 ; APPLICANT: OsteYoung, Katherine W.
 ; APPLICANT: Vitsha, Stanislav
 ; APPLICANT: Kokeharova, Olga A.
 ; APPLICANT: Gao, Honggo
 ; TITLE OF INVENTION: Plasmid Division and Related Genes and Proteins, and Methods of
 ; FILE REFERENCE: MSU-08153
 ; CURRENT APPLICATION NUMBER: US/10/600,070
 ; NUMBER OF SEQ ID NOS: 206
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 132
 ; LENGTH: 561
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (127)..(127)
 ; OTHER INFORMATION: n is a, c, g, or t
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (520)..(520)
 ; OTHER INFORMATION: n is a, c, g, or t
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (541)..(541)
 ; OTHER INFORMATION: n is a, c, g, or t
 ; US-10-600-070-132
 Query Match 11.0%; Score 402.2; DB 7; Length 561;
 Best Local Similarity 82.2%; Pred. No. 6.1e-95;
 Matches 514; Conservative 0; Mismatches 20; Indels 91; Gaps 1;
 Qy 2970 TTTCTTCTGTGATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3029
 Db 534 TTTCTTCTGTGATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 475
 Qy 3030 CACTTCCAGAAATGATGCTGTGAGCTGTGAGAGATATGATTCAGAGTGTGAGAAATTA 3089
 Db 474 CACTTCCAGAAATGATGCTGTGAGCTGTGAGAGATATGATTCAGAGTGTGAGAAATTA 415
 Qy 3090 AGTCTGTGCTTTTGGGCTGTATCATCCGATGAAATTTTCCAGAGGTGAGGAATAA 3149
 Db 414 AGTCTGTGCTTTTGGGCTGTATCATCCGATGAAATTTTCCAGAGGTGAGGAATAA 370
 Qy 3150 TCTACATTCATCAATGATGTGTGAAAATGTGTGACATGATTAATGCTGTGTGTGTGT 3209
 Db 369 ----- 370
 Qy 3210 TGATTTCTGTATTTATAGTTTGT 3269
 Db 369 ----- 370
 Qy 3270 GCTGAACTGTGAGAGCTGT 3329
 Db 325 GCTGAACTGTGAGAGCTGT 266
 Qy 3330 AGTGTGACATCTGACAGATGAAACCGGTGCTGTGTGTGAAAGCACTGTGAGAGATCT 3389

```

RESULT 9
US-10-600-070-184
; Sequence 184, Application US/10600070
; Publication No. US20040139500A1
; GENERAL INFORMATION:
; APPLICANT: Oosteryoung, Katherine W.
; APPLICANT: Vitsha, Stanislaw
; APPLICANT: Koksharova, Olga A.
; APPLICANT: Gao, Hongbo
; TITLE OF INVENTION: Placitid Division and Related Genes and Proteins, and Methods of
; TITLE OF INVENTION: Use
; FILE REFERENCE: MSU-08153
; CURRENT APPLICATION NUMBER: US/10/600,070
; CURRENT FILING DATE: 2003-06-20
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 184
; LENGTH: 631
; TYPE: DNA
; ORGANISM: Prunus persica
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (21)..(21)
; OTHER INFORMATION: n is a, c, g, or t
US-10-600-070-184

```

QY	2068	AATTCAATGCTGATG-----ACAAATGATGATCCCTCGGACTATGGAAATTG	2115
Db	424	AACTCAAGGATGACGATGACAAATGACAAATGACAAATGATCTTCTGACCTTTGCAAGCTA	483
QY	2116	TTGGAAACCTGTTGGCAGGGGTGTCTTCTTCTAGTTTCAGAGACACCAAGATATAAAAA	2175
Db	484	TTGGAGACGATGTTGATGAGAGTGGATTTCCCGAGTTTATGAGACACCAAGACATAGAG	543
QY	2176	TTTAACTCGGGGACTATGATGATGATCTTATGTTTGTAGTTACTTGGAAAGAGTGGAG	2235
Db	544	TTGAGACTGGGAGACTATGATGATGATCCCAAGCTTTGAGATATTGAAAGAGGCTGGAT	603
QY	2236	GTAGTTGAGGTTCTCCTTTAGCGCTG	2263
Db	604	GGCACATATGATTCACCTTAGCTGCTG	631

	Query Match	7.2%: Score 262.4; DB: 7; Length 1146;
	Best Local Similarity	Pred. No. 7.8e-58;
	Matches 630; Conservative 0; Mismatches 266; Indels 178; Gaps 3;	
QY	587 TCTGCTCCGCACGAAATGGGACCGACCGCTTCTCTCCGACTTCAATTGACACTCCGATT	646
Db	236 TATCCGCACACGAAATGGGACCGAGCGCTCAATGCGAATTTCAATTCCTCGGCACG	295
QY	647 CTTCTCTCTCTCTCTCTTGGCACCGGCCAACACACCGGCACTCTGCTTCTCTGACCAAT	706
Db	296 CCGCGCGCTTCACCTT-----CAACCTCCACTTCAGACCCCTCTCTCGCTCTCC	346
QY	707 CTATTGATCGTCCGGAAGCGACAGTCCCCCATTCGCATTGATTTCTACAGGTAATTAGAG	766
Db	347 GCCTCAATCTCTCGAGCGCTACGTGTGCATCTCTCGACCTGTACCGCAATTCGGCG	406
QY	767 CTCAACACAAATTTCTTAACCGATGGAATCAGAGAGCAATTCGAAAGCTAAGGTTTCGAAC	826
Db	407 CCGAACCAGCAATTCCTCGGAGACGGCAATTCGACGGGCTTACGAGGCCAAGTTCGAGAC	466
QY	827 GCGCGCAATTCGGTTTCAGGACGAGCGCTTTAATTCACGCGGACAGATTCCTCAACTG	886
Db	467 CTCTCTCAGTACGCTTTAGCAACGACTTTTAATCAGCCCGCCCAATTCCTCGACAG	526
QY	887 CTTCGGAACCTGTGTCTAATCCTCGGCTTAGAAGAGATCAATGAAGGTCCTTCTGATG	946
Db	527 CTGCGGAACCTTAGCTGATCTTAATCTTCGAGAGAGATCAATCAAGCCTTGTGACG	586
QY	947 ATGAAGAAGCTACAGTATCACTGATGTCTCTTGGAATTAAGTAAATTCGATTCGGAAAT	1006
Db	587 ACGAAGAAGCGCATTTCTACCTCAATATCCCTT-----	619


```

QY      1007 AATAAATTTCTCTGTTTAATTCAATGATGATTAAGAAGAACTTTATCTAGTG 1066
Db      620 -----TCGAC 624
QY      1067 AAGGTTCTGAGGCTCTCTGTGATTTGCAAGAAAGTGTTGAGACTGAGATGTTCTTCGG 1126
Db      625 AAGATTCCTGAGCGTTGTGTGTGTCAGAAAGACTGAGAGACGAGCTTGTGCTTGAG 684
QY      1127 GTTGGTAGGCTCGCTGCTTAAGAGAGGTTGCGCTTAAGTCGTTTAAGAAAGTGTGTTTA 1186
Db      685 ATTGGGCAAGGTTTCTCTTAAGAGAGGTTGCCAAGACGTTTAAGCAAGATGTGTGTG 744
QY      1187 GTTAAGCGCTTGCGTTCTCGATGTCCTCGAGGATGCTATGACATTGATCCACCTGAT 1246
Db      745 GCTATGCACTCGCATTTGTTGACGTGTCAAGGGAGTCTATGCTTGTCTCACCGGAT 804
QY      1247 TTTATTACTGTGATATGAGTTGTTGAGAAAGCTTTGAAGCTTTTAACAGTATTTGACTT 1306
Db      805 TTCAATGCGGCTGTGTGATGCTCGAAGGGCGTTGAAAGCTTTTGTGACGAA----- 855
QY      1307 GCTTGTGTAATTGACGAGCGTTGCTTTATTAAGAACTTTCTGATTTGATCTTTGTTA 1366
Db      856 ----- 855
QY      1367 TTGAGTCTGTGTAGAGAAAGAGCAAGTAGCCTTGCACCGGATTTACGTGCACAAATT 1426
Db      856 -----GAAAGGGCAACACGAGCTTGCATTCAGATTTTACAGACAAATA 897
QY      1427 GATGAGACTTTGGAAGGATCACTCCGCGTTATGCTTGAAGCTACTTGGCTTACCGCTT 1486
Db      898 GATGAGACGCTAGAGAGATPACCCACGTTGTGTTTGAACCTTTTACCTTGGCTCTT 957
QY      1487 GGTGATGATTAACGCTGCGAAGAACTAATATGTTTAAGCGTGTGCGAATAATTTGTGCG 1546
Db      958 GATGACGAACATCGAGCGCGGAGGAGAAAGTCTTCTGTGTGCTGAACATTTTGTGCG 1017
QY      1547 TCTGTTGAGAGAGTGTGAGCATCAGCTCTTGTGTTGGGGGTTTGAACCGTGAAGTTTATG 1606
Db      1018 GCGGTTGTGTGAGGGGGTGTGACAAATTTGCCCGGGGTTTCCACCGTGAAGCTTCATG 1077
QY      1607 AATGAGGCGTTTTTACGAATGACAGCTGTGAGCAGGTATTAACGTTTATGATATCC 1660
Db      1078 AATGAGGCAATCTTACACATGACAGCGGCTGAACAGGTGAACTTTTGTAGCC 1131

RESULT 11
US-10-600-070-135
; Sequence 135, Application US/10600070
; Publication No. US20040139500A1
GENERAL INFORMATION:
; APPLICANT: Osteoryoung, Katherine W.
; APPLICANT: Vittha, Stanislaw
; APPLICANT: Koksharova, Olga A.
; APPLICANT: Gao, Hongbo
; TITLE OF INVENTION: Placitid Division and Related Genes and Proteins, and Methods of
; FILE REFERENCE: MSU-08153
; CURRENT APPLICATION NUMBER: US/10/600,070
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 135
; LENGTH: 660
; TYPE: DNA
; ORGANISM: Medicago truncatula
; US-10-600-070-135

```

[illegible]

```

RESULT 12
US-10-600-070-126
; Sequence 126, Application US/10600070
; Publication No. US20040139500A1
GENERAL INFORMATION:
; APPLICANT: Osteeryoung, Katherine W.
; APPLICANT: Vilcha, Stanislav
; APPLICANT: Koksharova, Olga A.
; APPLICANT: Gao, Hongbo
; TITLE OF INVENTION: Placid Division and Related Genes and Proteins, and Methods of
; TITLE OR INVENTION: Use
; FILE REFERENCE: MSU-08153
; CURRENT APPLICATION NUMBER: US/10/600,070
; CURRENT FILING DATE: 2003-06-20
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 126
; LENGTH: 2283
; TYPE: DNA
; ORGANISM: Oryza sativa
US-10-600-070-126

```

	Query Match	5.7%	Score 208.8	DB 7	Length 660
	Best Local Similarity	60.3%	Pred. No. 7.8e-44		
	Matches 429	Conservative 0	Mismatches 247	Indels 76	Gaps 2
QY	561	ACGAGGCACAACCTTACATCATCTGCTCCGCCAGCAAAATGGCGCGACCGCTTCT			620

Query Match	5.3%	Score 194;	DB 7;	Length 2283;
Best Local Similarity	55.6%	Pred. No. 1.4e-39;		
Matches 514; Conservative	0;	Mismatches 300;	Indels 110;	Gaps 3;

OTHER INFORMATION: Clone ID: LIB3493-032-PL-M1-F8
US-09-732-627A-2154

Query Match 4.8%; Score 175; DB 3; Length 439;
Best Local Similarity 65.0%; Pred. No. 5e-35;
Matches 278; Conservative 0; Mismatches 141; Indels 9; Gaps 1;

```
QY 499 GTGGGATGTCCTCTCCCAATTCATTAATGCGGATTAACACCGGCGAGCAAAAGCTC 558
DB 12 GTTAGATCGGGCTGTGACCTCCGACACTGCTCTTATTGCAACCTGTAAACCTCTC 71
QY 559 CGACGATGACACACCTCTCAACTATCTGCTCGGCGAAGATGGGCGGACGCTT 618
DB 72 AATCTCCACACCGTTCACTACCGCTGCTGCTCGGCAATGAATGGGCGGACGCTC 131
QY 619 CTCTCCGACTCAATTTCACTCCG-----ATTCTCCCTCTCTCTCTGCGAC 669
DB 132 CTGTGACTTCCAAATTTCTCTCTGCCCCGATTAATTCGCTCTCTCTCTCTCTCA 191
QY 670 GCGACCAACGCGCACTCTGCTCTGCGACCACTTAATGATGTCGCGAAGCGCAC 729
DB 192 ACCGCACTCTTCTCCCTTAACCTCCCTTCTGCGCCCTTCTCTCTCGAAGCGCAC 251
QY 730 GTCCCATATCCCATGATTTTCAACAGGATTAAGAGCTCAACACATTTCTTAACGAT 789
DB 252 GTTTCATTTCCCTTGATTTTCAAGGTTTGAAGCGGACGACATTTCTTAAGTAT 311
QY 790 GGAATCAGAAAGACATTCGAACTAGGGTTTGAACCGCGCAATTCGGTTTCAAGCAG 849
DB 312 GGAATTAAGAAAGCTTATTAAGACGAGGTTTGAACCGCTCATATAGGTTCAAGTCA 371
QY 850 GACGCTTTAATGACCGGAGACAGATTTCTTAAGCTGCTTGGAAACTCTGTATATCT 909
DB 372 GACACATTAATTAAGCGGAAGACAGATTTCTTAAGCTGCTTGAAGCCTTAACCT 431
QY 910 CGGTCTAG 917
DB 432 GGGCTTAG 439
```

RESULT 15
US-10-600-070-175

Sequence 175, Application US/10600070
Publication No. US20040139500A1
GENERAL INFORMATION:
APPLICANT: Oeteryoung, Katherine W.
APPLICANT: Vitka, Stanislaw
APPLICANT: Kokeharova, Olga A.
APPLICANT: Gao, Hong
TITLE OF INVENTION: Plasmid Division and Related Genes and Proteins, and Methods of
FILE REFERENCE: MSU-08153
CURRENT APPLICATION NUMBER: US/10/600,070
CURRENT FILING DATE: 2003-06-20
NUMBER OF SEQ ID NOS: 206
SOFTWARE: PatentIn version 3.2
SEQ ID NO 175
LENGTH: 545
TYPE: DNA
ORGANISM: Gossypium arboreum
FEATURE:
NAME/KEY: misc_feature
LOCATION: (528)..(528)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (536)..(536)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (540)..(540)
OTHER INFORMATION: n is a, c, g, or t
US-10-600-070-175

Query Match 4.4%; Score 161; DB 7; Length 545;
Best Local Similarity 69.6%; Pred. No. 2.9e-31;
Matches 218; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

```
QY 3224 ATAGGTTTGGATGGGCGAATGCTTAAGATTGGACTGACAGACGACGTAAGCTGGCA 3283
DB 82 AGAGGTTCTGATGATGCTCAAAATGTTGAAGACATGACAGATCGTGAAGCGGAATGCTCA 141
QY 3284 GCTTGGTTGGTTATGATTAATACCTGTTGAACCTATCTGTTGACAGTGTGACAGTCTC 3343
DB 142 GCTTGGTTGGTTATGATTAATGCTTACTGAACATGCGCACTTGAAGTGTACCTGTTAC 201
QY 3344 AGCAGATGAAACCCGTGCTGTGTGAAGCACTCTGAGAGAGTCTGCTGTATCTGA 3403
DB 202 ACTAGATGGCGACGAGGCTGTAGTGAAGTACTCTGGAAGATCACTGCTGTTGACTGA 261
QY 3404 TTTGTTTATTCAGAAAACATGCTAGTGTGAGAACTTACACACAAAGATGCAAGT 3463
DB 262 TGTTCATCATCCGGAACAAATGCTTAAATGTAACTCTAACACAGAGATGAGAT 321
QY 3464 TTTCTGTCCAAGTCAAGGATGGAATAATCACTGAAGGCTCTGTTCTGATCATATATATAC 3523
DB 322 GTCTTGTTCAACTGAGGCTGGAATATCACTGAAGATCTGTCTAAGAACTTTAATCTATG 381
QY 3524 TCATATGTAGCAT 3536
DB 382 ATGTATTAAGCAT 394
```

Search completed: December 10, 2005, 18:16:03
Job time : 2829.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 10, 2005, 03:43:47 ; Search time 611.67 Seconds
(without alignments)
10656.596 Million cell updates/sec

Title: US-10-600-070B-3

Perfect score: 3667
Sequence: 1 Ggtctgcatgaagagaat.....ctatacataagggctacaa 3667

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 segs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents.NA:*
1: /cgn2_6/prodata/1/ina/1/COMB.seq:*
2: /cgn2_6/prodata/1/ina/5/COMB.seq:*
3: /cgn2_6/prodata/1/ina/6A/COMB.seq:*
4: /cgn2_6/prodata/1/ina/6B/COMB.seq:*
5: /cgn2_6/prodata/1/ina/H/COMB.seq:*
6: /cgn2_6/prodata/1/ina/PCUS/COMB.seq:*
7: /cgn2_6/prodata/1/ina/PP/COMB.seq:*
8: /cgn2_6/prodata/1/ina/RB/COMB.seq:*
9: /cgn2_6/prodata/1/ina/Backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	57.6	1.6	7218	2	US-08-232-463-14 Sequence 14, Appl
2	54	1.5	1141	3	US-09-806-708B-22 Sequence 22, Appl
3	50.2	1.4	57280	3	US-09-949-016-11796 Sequence 11796, A
4	50.2	1.4	57280	3	US-09-949-016-12843 Sequence 12843, A
5	50.2	1.4	57280	3	US-09-949-016-12844 Sequence 12844, A
6	50.2	1.4	57280	3	US-09-949-016-12846 Sequence 12846, A
7	50.2	1.4	57280	3	US-09-949-016-13542 Sequence 13542, A
8	50.2	1.4	57280	3	US-09-949-016-13543 Sequence 13543, A
9	50.2	1.4	57280	3	US-09-949-016-13544 Sequence 13544, A
10	50.2	1.4	57280	3	US-09-949-016-13545 Sequence 13545, A
11	50.2	1.4	57280	3	US-09-949-016-14633 Sequence 14633, A
12	50.2	1.4	57280	3	US-09-949-016-14634 Sequence 14634, A
13	50.2	1.4	57280	3	US-09-949-016-14635 Sequence 14635, A
14	50.2	1.4	57280	3	US-09-949-016-14636 Sequence 14636, A
15	50.2	1.4	57280	3	US-09-949-016-14637 Sequence 14637, A
16	50.2	1.4	57280	3	US-09-949-016-14638 Sequence 14638, A
17	50.2	1.4	57280	3	US-09-949-016-14639 Sequence 14639, A
18	50.2	1.4	57280	3	US-09-949-016-14640 Sequence 14640, A
19	49.4	1.3	913	2	US-08-217-327-3 Sequence 3, Appl
20	49.4	1.3	913	2	US-07-885-970A-3 Sequence 3, Appl
21	49.4	1.3	913	2	US-08-298-687A-3 Sequence 3, Appl
22	49.4	1.3	913	2	US-08-530-797-2 Sequence 2, Appl
23	49.4	1.3	913	2	US-08-298-829-3 Sequence 3, Appl
24	49.4	1.3	913	2	US-08-787-335-2 Sequence 2, Appl

25	49.4	1.3	1984	2	US-07-885-970A-25 Sequence 25, Appl
26	49.4	1.3	1985	2	US-08-298-687A-25 Sequence 25, Appl
27	49.4	1.3	1985	2	US-08-298-829-25 Sequence 25, Appl
28	49.2	1.3	324	3	US-09-547-693-234 Sequence 234, Appl
29	48.4	1.3	152132	3	US-09-949-016-13845 Sequence 12371, A
30	48.4	1.3	152145	3	US-09-949-016-12371 Sequence 12371, A
31	47.4	1.3	1141	3	US-10-148-806-3 Sequence 3, Appl
32	46.8	1.3	114793	3	US-09-949-016-14011 Sequence 14011, A
33	45	1.2	49818	3	US-09-949-016-12458 Sequence 12458, A
34	45	1.2	49829	3	US-09-949-016-14011 Sequence 14011, A
35	44.2	1.2	4403765	3	US-09-103-840A-2 Sequence 2, Appl
36	44.2	1.2	441529	3	US-09-103-840A-1 Sequence 2813, Ap
37	43.8	1.2	832	3	US-09-621-976-2813 Sequence 2813, Ap
38	43.8	1.2	832	3	US-09-621-976-2813 Sequence 23, Appl
39	42.8	1.2	1055	3	US-09-806-708B-23 Sequence 14080, A
40	42.8	1.2	60990	3	US-09-949-016-13813 Sequence 13813, A
41	42.4	1.2	23849	3	US-09-949-016-12384 Sequence 12384, A
42	42.4	1.2	39154	3	US-09-949-016-12801 Sequence 12801, A
43	42.4	1.2	39154	3	US-09-949-016-14326 Sequence 14326, A
44	42.4	1.2	39443	3	US-09-949-016-14326 Sequence 14326, A
45	42.4	1.2	39443	3	US-09-949-016-14327 Sequence 14327, A

ALIGNMENTS

RESULT 1
US-08-232-463-14
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHIEFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOXLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESS: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZspT-Fls

Best Local Similarity 53.9%; Pred. No. 0.033;
Matches 103; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

```
OY 565 AGCCACAACACCTCTACAACTATCTGCTCCGAGCAATGGGCGACCGCTCTCTCC 624
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 13101 AGCAGCATACACCAACCACTGCTCTCAACCATCAACCTTACCACTTATGAGCTCC 13042
OY 625 GACTTCATTTTCACTCCGATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 684
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 13041 ACCATCAATATCTCTATCACTGCTCAACCAACCGTCTTCACTATCAACCAAC 12982
OY 685 ACTCTGCTCTCTGCGCAACATCTATGATGCTGCGCAACGCGACGTCCCATCCCAT 744
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 12981 ACAACCAACCACTCATATCATGACTTCCACATCAACATTCATCATCACTCCACA 12922
OY 745 GATTCTACCA 755
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 12921 CATCATCAACA 12911
```

RESULT 4
US-09-949-016-12843/c

; Sequence 12843, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: CLO01307

; CURRENT APPLICATION NUMBER: US/09/949,016

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 12843

; TYPE: DNA

; ORGANISM: Human

; US-09-949-016-12843

Query Match 1.4%; Score 50.2; DB 3; Length 57280;

Best Local Similarity 53.9%; Pred. No. 0.033;
Matches 103; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

```
OY 565 AGCCACAACACCTCTACAACTATCTGCTCCGAGCAATGGGCGACCGCTCTCTCC 624
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 13101 AGCAGCATACACCAACCACTGCTCTCAACCATCAACCTTACCACTTATGAGCTCC 13042
OY 625 GACTTCATTTTCACTCCGATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 684
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 13041 ACCATCAATATCTCTATCACTGCTCAACCAACCGTCTTCACTATCAACCAAC 12982
OY 685 ACTCTGCTCTCTGCGCAACATCTATGATGCTGCGCAACGCGACGTCCCATCCCAT 744
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 12981 ACAACCAACCACTCATATCATGACTTCCACATCAACATTCATCATCACTCCACA 12922
OY 745 GATTCTACCA 755
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 12921 CATCATCAACA 12911
```

RESULT 5
US-09-949-016-12844/c

; Sequence 12844, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CLO01307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 12844

; LENGTH: 57280

; TYPE: DNA

; ORGANISM: Human

; US-09-949-016-12844

Query Match 1.4%; Score 50.2; DB 3; Length 57280;

Best Local Similarity 53.9%; Pred. No. 0.033;
Matches 103; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

```
OY 565 AGCCACAACACCTCTACAACTATCTGCTCCGAGCAATGGGCGACCGCTCTCTCC 624
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 13101 AGCAGCATACACCAACCACTGCTCTCAACCATCAACCTTACCACTTATGAGCTCC 13042
OY 625 GACTTCATTTTCACTCCGATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 684
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 13041 ACCATCAATATCTCTATCACTGCTCAACCAACCGTCTTCACTATCAACCAAC 12982
OY 685 ACTCTGCTCTCTGCGCAACATCTATGATGCTGCGCAACGCGACGTCCCATCCCAT 744
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 12981 ACAACCAACCACTCATATCATGACTTCCACATCAACATTCATCATCACTCCACA 12922
OY 745 GATTCTACCA 755
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 12921 CATCATCAACA 12911
```

RESULT 6
US-09-949-016-12846/c

; Sequence 12846, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CLO01307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 12846

; LENGTH: 57280

; TYPE: DNA

; ORGANISM: Human

; US-09-949-016-12846

Query Match 1.4%; Score 50.2; DB 3; Length 57280;

Best Local Similarity 53.9%; Pred. No. 0.033;
Matches 103; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

```
OY 565 AGCCACAACACCTCTACAACTATCTGCTCCGAGCAATGGGCGACCGCTCTCTCC 624
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 13101 AGCAGCATACACCAACCACTGCTCTCAACCATCAACCTTACCACTTATGAGCTCC 13042
OY 625 GACTTCATTTTCACTCCGATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 684
```

Db 13041 ACCATCAATATCCCTATACAGTCCACCAACCGCTTTCCACCATCACCACCAACC 12982
QY 685 ACTCTGCTCTCTGCGACCAATCTATGATGTCGCGAAGCCAGCGCCCATCCCAT 744
Db 12981 ACAACCAACCACTCACTATCATGACTTCCACCATTAACATTAATCATCTCCACCAAC 12922
QY 745 GATTTCTACCA 755
Db 12921 CATCATCAGCA 12911

RESULT 7
US-09-949-016-13542/c
; Sequence 13542, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 13542
; LENGTH: 57280
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13542

Query Match 1.4%; Score 50.2; DB 3; Length 57280;
Best Local Similarity 53.9%; Pred. No. 0.033;
Matches 103; Conservative 0; Mismatches 88; Indels 0; Gaps 0;
QY 565 AGCCACAACCTCTTACATATCTGCTCCGCGCAAAATGGCCGACCGCTTCTCTCC 624
Db 13101 AGCACCAATCAACCAACCAACCGCTTCCACCATTAATGAGCTCC 13042
QY 625 GACTCAATTTCACTCCGATTCCTCCCTCCCTTGGCCAGCGACACCGCC 684
Db 13041 ACATCAATATCCCTATACAGTCCACCAACCGCTTCCACCATTAATGAGCTCC 12982
QY 685 ACTCTGCTCTCTGCGACCAATCTATGATGTCGCGAAGCCAGCGCCCATCCCAT 744
Db 12981 ACAACCAACCACTCACTATCATGACTTCCACCATTAACATTAATCATCTCCACCAAC 12922
QY 745 GATTTCTACCA 755
Db 12921 CATCATCAGCA 12911

RESULT 8
US-09-949-016-13543/c
; Sequence 13543, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 13543
; LENGTH: 57280
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13543

Query Match 1.4%; Score 50.2; DB 3; Length 57280;
Best Local Similarity 53.9%; Pred. No. 0.033;
Matches 103; Conservative 0; Mismatches 88; Indels 0; Gaps 0;
QY 565 AGCCACAACCTCTTACATATCTGCTCCGCGCAAAATGGCCGACCGCTTCTCTCC 624
Db 13101 AGCACCAATCAACCAACCAACCGCTTCCACCATTAATGAGCTCC 13042
QY 625 GACTCAATTTCACTCCGATTCCTCCCTCCCTTGGCCAGCGCCACACCGCC 684
Db 13041 ACATCAATATCCCTATACAGTCCACCAACCGCTTCCACCATTAATGAGCTCC 12982
QY 685 ACTCTGCTCTCTGCGACCAATCTATGATGTCGCGAAGCCAGCGCCCATCCCAT 744
Db 12981 ACAACCAACCACTCACTATCATGACTTCCACCATTAACATTAATCATCTCCACCAAC 12922
QY 745 GATTTCTACCA 755
Db 12921 CATCATCAGCA 12911

RESULT 9
US-09-949-016-13544/c
; Sequence 13544, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 13544
; LENGTH: 57280
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13544

Query Match 1.4%; Score 50.2; DB 3; Length 57280;
Best Local Similarity 53.9%; Pred. No. 0.033;
Matches 103; Conservative 0; Mismatches 88; Indels 0; Gaps 0;
QY 565 AGCCACAACCTCTTACATATCTGCTCCGCGCAAAATGGCCGACCGCTTCTCTCC 624
Db 13101 AGCACCAATCAACCAACCAACCGCTTCCACCATTAATGAGCTCC 13042
QY 625 GACTCAATTTCACTCCGATTCCTCCCTCCCTTGGCCAGCGCCAGCGCC 684
Db 13041 ACATCAATATCCCTATACAGTCCACCAACCGCTTCCACCATTAATGAGCTCC 12982
QY 685 ACTCTGCTCTCTGCGACCAATCTATGATGTCGCGAAGCCAGCGCCCATCCCAT 744
Db 12981 ACAACCAACCACTCACTATCATGACTTCCACCATTAACATTAATCATCTCCACCAAC 12922
QY 745 GATTTCTACCA 755

GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CLO01307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 14635
LENGTH: 57280
TYPE: DNA
ORGANISM: Human
US-09-949-016-14635

Query Match 1.4%; Score 50.2; DB 3; Length 57280;
Best Local Similarity 53.9%; Pred. No. 0.033; Mismatches 88; Indels 0; Gaps 0;
Matches 103; Conservative 0;

QY 565 AGCCACAACCTCTACACTATCTGCTCCGCGACGAATGGCGGACCGTCTTCTCTCC 624
DB 13101 AGCACCATCACCAACCAATGCGCTCCACACCATCACACCTTACCATATGAGCTCC 13042
QY 625 GACTTCAATTTCACTCCGATTCCTCTCTCTCTTCCGACCGCACCAACCAACCGCC 684
DB 13041 ACCATCAATATCTCCATATACCGTCCACCAACCGCTTACCATCACCAACCAACCG 12982
QY 685 ACTCTGCTCTCTGCGACCATCTATGATGTCGCCGACGACGATCCCATCCCATTT 744
DB 12981 ACAACCACTCACTACTATGACTTCCACCATCAACTTACCATCACTCCACCAAC 12922
QY 745 GATTTCTACCA 755
DB 12921 CATCATCACA 12911

RESULT 14
US-09-949-016-14636/c
Sequence 14636, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CLO01307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 14636
LENGTH: 57280
TYPE: DNA
ORGANISM: Human
US-09-949-016-14636

Query Match 1.4%; Score 50.2; DB 3; Length 57280;
Best Local Similarity 53.9%; Pred. No. 0.033; Mismatches 88; Indels 0; Gaps 0;
Matches 103; Conservative 0;

QY 565 AGCCACAACCTCTACACTATCTGCTCCGCGACGAATGGCGGACCGTCTTCTCTCC 624
DB 13101 AGCACCATCACCAACCAATGCGCTCCACACCATCACACCTTACCATATGAGCTCC 13042

DB 13101 AGCACCATCACCAACCAACTGCTCCACCAACATCACCACTTACCATATGAGCTCC 13042
QY 625 GACTTCAATTTCACTCCGATTCCTCTCTCTCTTCCGACCGCACCAACCAACCGCC 684
DB 13041 ACCATCAATATCTCCATATACCGTCCACCAACCGCTTACCATCACCAACCAACCG 12982
QY 685 ACTCTGCTCTCTGCGACCATCTATGATGTCGCCGACGACGATCCCATCCCATTT 744
DB 12981 ACAACCACTCACTACTATGACTTCCACCATCAACTTACCATCACTCCACCAAC 12922
QY 745 GATTTCTACCA 755
DB 12921 CATCATCACA 12911

RESULT 15
US-09-949-016-14637/c
Sequence 14637, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CLO01307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 14637
LENGTH: 57280
TYPE: DNA
ORGANISM: Human
US-09-949-016-14637

Query Match 1.4%; Score 50.2; DB 3; Length 57280;
Best Local Similarity 53.9%; Pred. No. 0.033; Mismatches 88; Indels 0; Gaps 0;
Matches 103; Conservative 0;

QY 565 AGCCACAACCTCTACACTATCTGCTCCGCGACGAATGGCGGACCGTCTTCTCTCC 624
DB 13101 AGCACCATCACCAACCAATGCGCTCCACACCATCACACCTTACCATATGAGCTCC 13042
QY 625 GACTTCAATTTCACTCCGATTCCTCTCTCTCTTCCGACCGCACCAACCAACCGCC 684
DB 13041 ACCATCAATATCTCCATATACCGTCCACCAACCGCTTACCATCACCAACCAACCG 12982
QY 685 ACTCTGCTCTCTGCGACCATCTATGATGTCGCCGACGACGATCCCATCCCATTT 744
DB 12981 ACAACCACTCACTACTATGACTTCCACCATCAACTTACCATCACTCCACCAAC 12922
QY 745 GATTTCTACCA 755
DB 12921 CATCATCACA 12911

Search completed: December 10, 2005, 08:09:24
Job time : 617.67 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 10, 2005, 16:58:05 ; Search time 13948.2 Seconds
(without alignments)
12300.344 Million cell updates/sec

Title: US-10-600-070B-3

Perfect score: 3667
Sequence: 1 tgcctgcataaagagagaat.....ctataacataagggctacaa 3667

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST:
1: gb_est1:
2: gb_est2:
3: gb_est3:
4: gb_hic:
5: gb_est4:
6: gb_est5:
7: gb_est6:
8: gb_est7:
9: gb_gsa1:
10: gb_gsa2:
11: gb_gsa3:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	ID	Description
1	725.2	19.8	919	4	CNS09YJH
2	436.2	11.9	741	4	CNS09ZSS
3	434.6	11.9	741	4	CNS09JNM
4	402.2	11.0	561	1	AI998415
5	376.4	10.3	415	3	BP782886
6	363.6	9.9	785	3	BP785511
7	359.2	9.8	434	3	BP620404
8	359.2	9.8	434	3	BP620404
9	359.2	9.8	434	3	BP620404
10	359.2	9.8	434	3	BP620404
11	334	9.1	871	10	DU050213
12	315.8	8.6	534	1	AV830764
13	311.8	8.5	320	9	BZ765587
14	307.2	8.4	320	9	BZ765587
15	305.6	8.3	320	9	BZ765587
16	304.4	8.3	320	9	BZ765587
17	299.4	8.2	311	8	AY199896
18	295.6	8.1	631	5	BU046755
19	293.4	8.0	832	7	CO079829
20	265.4	7.2	751	11	CR486014
21	254	6.9	849	7	CO117047
22	252.6	6.9	769	7	CO075595

23	252	6.9	813	7	CO071968	CO071968 GR_Ea30N
24	244.4	6.7	741	8	CX543653	CX543653 UCRPT01.5
25	239.6	6.5	423	1	AV812946	AV812946
26	237.2	6.5	703	7	CN918842	CN918842 030207ZABP
27	226	6.2	728	8	CX674370	CX674370 UCRCS10.6
28	212.6	5.8	521	7	CN914655	CN914655 030110ABN
29	212.6	5.8	593	7	CN904734	CN904734 010926ABD
30	210	5.7	215	10	AL952467	AL952467 Arabidops
31	208.8	5.7	660	2	B1268376	B1268376 NF19C11
32	202.8	5.5	697	8	CX173275	CX173275 B01.69-11
33	194	5.3	2307	10	CF603268	CF603268 BACCA01.0
34	193.2	5.3	510	6	BO586205	BO586205 E012392-0
35	166.4	4.5	542	5	CV241483	CV241483 MS02512.B
36	161.8	4.4	819	7	CK090561	CK090561 F013P64.3
37	161.8	4.4	922	7	CV264112	CV264112 MS02023.B
38	161.8	4.4	545	5	BO410206	BO410206 GA_E0002
39	161	4.4	692	10	CL722297	CL722297 OR_BBA005
40	158.8	4.3	831	8	DN939692	DN939692 6602.3 Af
41	158.8	4.3	808	2	BE472035	BE472035 EST416888
42	157.8	4.3	722	7	CO117046	CO117046 GR_Eb019
43	157.4	4.3	321	10	AL936312	AL936312 Arabidops
44	156.2	4.3	947	10	CG088802	CG088802 PUBN47TB
45	156.2	4.3	947	10	CG088802	CG088802 PUBN47TB

ALIGNMENTS

RESULT 1
CNS09YJH 919 bp mRNA linear HTC 04-FEB-2004
LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone
DEFINITION GSLT89ZC08 of Adult vegetative tissue of strain col-0 of
Arabidopsis thaliana (thale cress).

ACCESSION BX841670 GI:42406830
VERSION BX841670.1
KEYWORDS HTC; GSLT cDNA.
SOURCES Arabidopsis thaliana
ORGANISM Arabidopsis thaliana (thale cress)

REFERENCE
AUTHORS Castellani, V., Aury, J.M., Jallion, O., Wincker, P., Clepet, C., Menard, M., Cruaud, C., Queirer, F., Scarpelli, C., Schachter, V., Temple, G., Caboche, M., Weissenbach, J. and Salanoubat, M.
Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation

TITLE
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 919)
AUTHORS Direct Submission
JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT The sequences are based on single pass reads. Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G. Genoscope members carried out sequencing and annotation : Castellani V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M. URGV INRA : Clepet C., Caboche M.

Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Pnap. http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full length
http://www.genoscope.cns.fr/cgi-bin/gsb/gsb?source=Arabidopsis

FEATURES
source location/qualifiers
1..919
/organism="Arabidopsis thaliana"
/mol_type="mRNA"

Matches	548;	Conservative	0;	Mismatches	18;	Indels	91;	Gaps	1;
Qy	2970	TTTCTTCCCTGATATTTTCTTTCTTTGATTAGGGTCAGTCAGAGCTGACGATTGCAAG		3029					
Db	176	TTTCTTCTATGATATGATGTCGACATACATAGGGTCAGTCAGAGCTGACGATTGCAAG		235					
Qy	3030	CACCTTCCCGAATGATGCTAGACATGCGAGAAATTTAGTATCCAAAGTGGCAGAGATT		3089					
Db	236	CACCTTCCCGAATGATGCTAGACATGCGAGAAATTTAGTATCCAAAGTGGCAGAGATT		295					
Qy	3090	AGTCTCTGGCTTTGGGCGCTGATACCGATGAAATGTTACGAGGAGGAAATAA		3149					
Db	296	AGTCTCTGGCTTTGGGCGCTGATACCGATGAAATGTTACGAGGAGGAAATAA		340					
Qy	3150	TCTACAAATTCATCAATGTTGTGAAAACCTGTCACATGATTATGTCGTGCTTGT		3209					
Db	341	-----		340					
Qy	3210	TGATTTCTGTTATTTATAGTTTGGATGGCGAATGCTGAAGATTGACAGAGCA		3269					
Db	341	-----AGGTTTGGATGGCGAATGCTGAAGATTGACAGAGCA		384					
Qy	3270	GCTGAAACGCGAGCTGGGTTGGTTATGATTATACCTGTTGAAACTATCTGTGAC		3329					
Db	385	GCTGAAACGCGAGCTGGGTTGGTTATGATTATACCTGTTGAAACTATCTGTGAC		444					
Qy	3330	AGTGTGACAGCTCTCAGCAGATGAAACCGTGTCTGTGTGGAAGCAACTGTGAGAGCT		3389					
Db	445	AGTGTGACAGCTCTCAGCAGATGAAACCGTGTCTGTGTGGAAGCAACTGTGAGAGCT		504					
Qy	3390	GCTTGTCTATCTGATTTGGTTTCATCCAGAAAACAATGCTACTGATGTCCAGAACT		3449					
Db	505	GCTTGTCTATCTGATTTGGTTTCATCCAGAAAACAATGCTACTGATGTCCAGAACT		564					
Qy	3450	ACAAGATGCGAGTTTCTGTGTCAGAGCTGGGAAAATGCACTGGAAGCTCTGTCT		3509					
Db	565	ACAAGATGCGAGTTTCTGTGTCAGAGCTGGGAAAATGCACTGGAAGCTCTGTCT		624					
Qy	3510	GCATCATATATATCTCATATGTAGACATGTCTGAGCTTGGAGATTCTCTTTGTTGTA		3569					
Db	625	GCATCATATATATCTCATATGTAGACATGTCTGAGCTTGGAGATTCTCTTTGTTGTA		684					
Qy	3570	ATTCTCTCTAAGTTAGTTGTTTAAATGAACAAAAAATTAAAGTTCTTGCA		3626					
Db	685	ATTCTCTCTAAGTTAGTTGTTTAAATGAACAAAAAATTAAAGTTCTTGCA		741					
RESULT 3									
CNS09YNM	741 bp	mRNA	linear	HTC 06-FEB-2004					
LOCUS	Arabidopsis thaliana Full-length cDNA Complete sequence from clone								
DEFINITION	GSLSI1592A05 of Siliques of strain col-0 of Arabidopsis thaliana								
	(thale cress).								
ACCESSION	BX83489								
VERSION	BX83489.1	GI:42455179							
KEYWORDS	HTC; GSIT cDNA.								
SOURCE	Arabidopsis thaliana (thale cress)								
ORGANISM	Arabidopsis thaliana								
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;								
AUTHORS	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;								
	Rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.								
	1 (bases 1 to 741)								
	Castelli V., Aury J.M., Jallion O., Winkler P., Clepet C.,								
	Menard M., Cruaud C., Quetier F., Scarpelli C., Schachter V.,								
	Temple G., Caboche M., Weissenbach J., and Salanoubat M.								
	Whole Genome Sequence Comparisons and Full-Length cDNA Sequences:								
	A Combined Approach to Evaluate and Improve Arabidopsis Genome								
	Annotation								
	Unpublished								
JOURNAL	2 (bases 1 to 741)								
REFERENCE	Genoscope.								
AUTHORS	Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :								
JOURNAL									

COMMENT									
BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr									
- Web : www.genoscope.cns.fr)									
The sequences are based on single pass reads.									
Life Technologies (a division of Invitrogen) members carried out									
full-length libraries construction; Temple G.									
Genoscope members carried out sequencing and annotation : Castelli									
V., Aury J.M., Jallion O., Winkler P., Clepet C.,									
Schachter V., Weissenbach J., Salanoubat M.									
URGV INRA : Clepet C., Caboche M.									
Annotation is based on the June 2003 version of the Arabidopsis									
genome released by MIPS (Munich Information center for Protein									
Sequences). 5 prime and 3 prime are assembled with Phrap.									
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/full									
length									
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.									
FEATURES									
source									
1..741									
/organism="Arabidopsis thaliana"									
/mol_type="mRNA"									
/db_xref="taxon:3702"									
/clone="GSLSI1592A05"									
/tissue_type="Siliques"									
/ecotype="Col-0"									
/plasmid="PCWSPORT_6"									
1..741									
/gene="At5g42480"									
ORIGIN									
Query Match 11.9%; Score 434.6; DB 4; Length 741;									
Best Local Similarity 83.3%; Pred. No. 1.1e-100;									
Matches 547; Conservative 0; Mismatches 19; Indels 91; Gaps 1;									
Qy	2970	TTTCTTCCCTGATATTTTCTTTCTTTGATTAGGGTCAGTCAGAGCTGACGATTGCAAG		3029					
Db	176	TTTCTTCTATGATATGATGTCGACATACATAGGGTCAGTCAGAGCTGACGATTGCAAG		235					
Qy	3030	CACCTTCCCGAATGATGCTAGACATGCGAGAAATTTAGTATCCAAAGTGGCAGAGATT		3089					
Db	236	CACCTTCCCGAATGATGCTAGACATGCGAGAAATTTAGTATCCAAAGTGGCAGAGATT		295					
Qy	3090	AGTCTCTGGCTTTGGGCGCTGATACCGATGAAATGTTACGAGGAGGAAATAA		3149					
Db	296	AGTCTCTGGCTTTGGGCGCTGATACCGATGAAATGTTACGAGGAGGAAATAA		340					
Qy	3150	TCTACAAATTCATCAATGTTGTGAAAACCTGTCACATGATTATGTCGTGCTTGT		3209					
Db	341	-----		340					
Qy	3210	TGATTTCTGTTATTTATAGTTTGGATGGCGAATGCTGAAGATTGACAGAGCA		3269					
Db	341	-----AGTTTTGGATGGCGAATGCTGAAGATTGACAGAGCA		384					
Qy	3270	GCTGAAACGCGAGCTGGGTTGGTTATGATTATACCTGTTGAAACTATCTGTGAC		3329					
Db	385	GCTGAAACGCGAGCTGGGTTGGTTATGATTATACCTGTTGAAACTATCTGTGAC		444					
Qy	3330	AGTGTGACAGCTCTCAGCAGATGAAACCGTGTCTGTGTGGAAGCAACTGTGAGAGCT		3389					
Db	445	AGTGTGACAGCTCTCAGCAGATGAAACCGTGTCTGTGTGGAAGCAACTGTGAGAGCT		504					
Qy	3390	GCTTGTCTATCTGATTTGGTTTCATCCAGAAAACAATGCTACTGATGTCCAGAACT		3449					
Db	505	GCTTGTCTATCTGATTTGGTTTCATCCAGAAAACAATGCTACTGATGTCCAGAACT		564					
Qy	3450	ACAAGATGCGAGTTTCTGTGTCAGAGCTGGGAAAATGCACTGGAAGCTCTGTCT		3509					
Db	565	ACAAGATGCGAGTTTCTGTGTCAGAGCTGGGAAAATGCACTGGAAGCTCTGTCT		624					
Qy	3510	GCATCATATATCTCATATGTAGACATGTCTGAGCTTGGAGATTCTCTTTGTTGTA		3569					
Db	625	GCATCATATATCTCATATGTAGACATGTCTGAGCTTGGAGATTCTCTTTGTTGTA		684					
Qy	3570	ATTCTCTCTAAGTTAGTTGTTTAAATGAACAAAAAATTAAAGTTCTTGCA		3626					

Db 685 ATTCTCTCTTAAGTTAGTTTAAATGAACCAAAAATTTAACTCTTGGCA 741

RESULT 4
 A1998415/c 561 bp mRNA linear EST 08-SEP-1999

LOCUS A1998415/c
 DEFINITION thaliana cDNA clone 701545606, rosette-2 Arabidopsis

ACCESSION A1998415
 VERSION A1998415.1 GI:5845320

KEYWORDS EST.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

REFERENCE 1
 Arabidopsis thaliana (Col-0, rosette-2 Arabidopsis. Eukaryote, Viridiplantae, Streptophyta, Embryophyta, Tracheophyta, Spermatophytes, Magnoliophyta, eudicotyledons, core eudicotyledons; rosids, eurosids II, Brassicales, Brassicaceae, Arabidopsis. 1 (bases 1 to 561)

AUTHORS Chen, D., Momiyama, M., Chan, E., Mooney, M., Cartoon, B., Gilliland, D., Wang, X., Hillman, J., Giegler, K., Kim, C., Doyle, M., Brzozka, P., Gorgone, G., Burns, D., Griffin, J., Mouanoutou, M., Nguyen, D., Tan, R., Rose, M., Warren, B., Ton, B., Kasuriy, K., Borillo, C., Cardio, T., Policky, J., Suzuki, G., Argentine, C., Shah, S., Nobrega, A., Murry, L., Turner, C., Krizorian, S., Elder, L. and Hanson, D.

TITLE Arabidopsis thaliana Gene Expression Microarray

JOURNAL Unpublished (1999)

COMMENT Contact: David Smoller, Ph.D.
 Genome Systems, Inc., a wholly owned subsidiary of Incyte Pharmaceuticals, Inc.
 4633 World Parkway Circle, St. Louis, MO 63134, USA
 Tel: 877-577-2733
 Fax: 314-427-3324
 Email: service@genomesystems.com.

FEATURES
 source Location/Qualifiers
 1..561
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /ecotype="Col-0"
 /db_xref="taxon:3702"
 /clone="701545606"
 /tissue_type="rosette"
 /dev_stage="4 - 7 weeks"
 /clone_lib="A. thaliana, Columbia Col-0, rosette-2"
 /note="Vector: pSPOR; Site 1: NotI, Site 2: SalI, cDNA library was derived from untreated rosette tissue from Arabidopsis thaliana, Columbia Col-0, at 4 - 7 weeks. Plants were grown in 1:1:1 peat moss/vermiculite/perlite soil at 22 deg. C +/- 3 deg. C under constant light, and watered with fertilizer. cDNA synthesis was initiated using a NotI-oligo(dt) primer. Double-stranded cDNA was bluntended, ligated to SalI adaptors, digested with NotI, size-selected, and cloned into the NotI and SalI sites of the pSPORT vector."

ORIGIN

Query Match 11.0%; Score 402.2; DB 1; Length 561;
 Best Local Similarity 82.2%; Pred. No. 2,4e-92;
 Matches 514; Conservative 0; Mismatches 20; Indels 91; Gaps 1;

QY 2970 TTCTTCCTCGATGATTTTTCCTTTTCATTTAGGGCTGTCAGAGCTGACATTCAGAG 3029
 Db 534 TTTCTTCTCATGGAACGATGATGCTCGTACCAATAGGGCTGATGAGCTGACGATTCAGAG 475

OY 3030 CACTTCCAGAAATGATGCTAGAGCTGACAGAGAAATATAGTATCCAGTGCAGAGATTA 3089
 Db 474 CACTTCCAGAAATGATGCTAGAGCTGACAGAGAAATATAGTATCCAGTGCAGAGATTA 415

OY 3090 AGTCTTGGCTTTTGGGCTGTGATACCGCATAGAAATGTTACCAAGGTAGGGAATAA 3149
 Db 414 AGTCTTGGCTTTTGGGCTGTGATACCGCATAGAAATGTTACCAAGGTAGGGAATAA 3149

OY 3150 TCTACAAATCAATCAATGATGCTGAAAACGTGTGAGCATGATTAATAGTCTGCTTGT 3209

QY	Db	369	-----	370
QY	3210	TGATTCGTGTTATTTATAGCTTTTGAATGGCGCAATGCTGAAAGATTGGACTGCAGAGCA		3269
Db	369	-----	AGTTTTGGAGTGGCGCAATGCTGAAGATTGGACTGCAGAGCA	326
QY	3270	GCTGAAACCTGCGCAGCTTGGGGTGGTTTATGATTATACCTGTTGAAACATATCTGTTGAC		3329
Db	325	GCTGAACCTGCGCAGCTTGGGGTGGTTTATGATTATACCTGTTGAAACATATCTGTTGAC		266
QY	3330	AGTGTGACAGTCTTCAGCAGATGGAACCCGTCCTCTGCTGGAAGCACTTCGAGAGTCT		3389
Db	265	AGTGTGACAGTCTTCAGCAGATGGAACCCGTCCTCTGCTGGAAGCACTTCGAGAGTCT		206
QY	3390	GCTTGCTATCTGATTTGGTTGCTTCCAGAAACAATCTATCTGATGTGACAACTTACACA		3449
Db	205	GCTTGCTATCTGATTTGGTTGCTTCCAGAAACAATCTATCTGATGTGACAACTTACACA		146
QY	3450	ACAAGATACGAAGTTTCTGCTCCAGTTCAGGGTGAAGAAATCACTGAAGCTCTGTTCTT		3509
Db	145	ACAAGATACGAAGTTTCTGCTCCAGTTCAGGGTGAAGAAATCACTGAAGCTCTGTTCTT		86
QY	3510	GCATCATATATATCTCATATGATGACATGCTGACCTTGCAAGATTCCTTTGTTCTGTAA		3569
Db	85	GCATCATATATATCTCATATGATGACATGCTGACCTTGCAAGATTCCTTTGTTGTTAA		26
QY	3570	ATTCTCTCTCAAGTTAGTGTAT	3594	
Db	25	ATTCTCTCTCAAGTTAGTGTAT	1	
RESULT 5				
BP782886/c				
LOCUS		410 bp	mRNA	linear
DEFINITION	BP782886	RAFL7	Arabidopsis thaliana	cDNA clone RAFL07-83-O20 3',
ACCESSION	BP782886			mRNA sequence.
VERSION	BP782886.1	GI:58842065		EST.
KEYWORDS				Arabidopsis thaliana (thale cress)
SOURCE				Arabidopsis thaliana
ORGANISM				Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE				1 (bases 1 to 410)
AUTHORS				Seki, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M., Akiyama, K., Iida, K., Enju, A., Sakurai, T., Arahawa, T., Caminici, P., Fukuda, S., Iida, J., Kawai, J., Sasaki, D., Shiraki, T., Hayashizaki, Y. and Shimozaki, K.
TITLE				Large-Scale Analysis of RIKEN Arabidopsis Full-length cDNAs
JOURNAL				Unpublished (2005)
COMMENT				Contact: Motoaki Seki Plant Functional Genomics Research Group RIKEN Genomic Sciences Center, 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan Tel: 81-298-36-4359 Fax: 81-298-36-9060 Email: meeki@rc.riken.go.jp An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998, 2002). This clone is in a modified pBluescript Vector. Please visit our web site (http://pfweb.gsc.riken.jp and http://large.gsc.riken.jp) for further details. reversed clone.
FEATURES				location/Qualifiers
source				1..410
				/organism="Arabidopsis thaliana"
				/mol_type="mRNA"
				/db_xref="taxon:3702"
				/clone="RAFL07-83-O20"
				/dev_stage="rossette plante"
				/lab_host="DH10B"

/clone.lib="RAFL7"
/note="Site 1: BamHI; site 2: SalI; subjected to
cold-treated (1, 2, 5, 10, 24 hr)"

ORIGIN

Query Match 10.3%; Score 376.4; DB 3; Length 410;
Best Local Similarity 99.5%; Pred. No. 1e-85;
Matches 388; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

3224 ATAGGTTTGGATGGCGAATGCTGAAGATTGGACTGACAGACAGCGTGAACCTGCCA 3283
389 AGAGGTTTGGATGGCGAATGCTGAAGATTGGACTGACAGACAGCGTGAACCTGCCA 330
3284 GCTTGGGTTTATGATTATATACACTGTGAACTATCTGTGACAGTGTGACAGTCTC 3343
329 GCTTGGGTTTATGATTATATACACTGTGAACTATCTGTGACAGTGTGACAGTCTC 270
3344 AGCAGATGAAACCCGTGCTGTGTGGAAGCAACTGTGAGAGTCTGCTGTCTATCTGA 3403
269 AGCAGATGAAACCCGTGCTGTGTGGAAGCAACTGTGAGAGTCTGCTGTCTATCTGA 210
3404 TTTGGTTTCATCCAGAAACAATGCTACTGATGTGAGAACTTACACAACAAGATGAGAGT 3463
209 TTTGGTTTCATCCAGAAACAATGCTACTGATGTGAGAACTTACACAACAAGATGAGAGT 150
3464 TTTCTGTCGAAGTCAGGGTGGAAATCACTGAAGCTCTGTTCTTGATCATATATATAC 3523
149 TTTCTGTCGAAGTCAGGGTGGAAATCACTGAAGCTCTGTTCTTGATCATATATATAC 90
3524 TCATATGATGATGTCTGAGCTTGGAGATTCTCTTGTCTGTAATTTCTCTCTAAG 3583
89 TCATATGATGATGTCTGAGCTTGGAGATTCTCTTGT- TGTAATTTCTCTCTAAG 31
3584 TTAGGTTTATTAATGAACAAAAAATT 3613
30 TTAGGTTTATTAATGAACAAAAAATT 1

RESULT 6

BP785511/c 415 bp mRNA linear EST 10-FEB-2005
LOCUS BP785511 RAFL7 Arabidopsis thaliana cDNA clone RAFL07-95-D11 3,
DEFINITION mRNA sequence.

ACCESSION BP785511
VERSION BP785511.1 GI:59261187
KEYWORDS EST.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

REFERENCE 1 (bases 1 to 415)
Seki,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M., Akiyama,K.,
Iida,K., Enju,A., Sakurai,T., Arakawa,T., Carninci,P., Fukuda,S.,
Iida,J., Kawai,D., Sasaki,D., Shiraki,T., Hayashizaki,Y. and
Shinozaki,K.
Large-Scale Analysis of RIKEN Arabidopsis Full-length cDNAs
Unpublished (2005)

AUTHORS

TITLE JOURNAL
COMMENT Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: msekic@rct.riken.go.jp
An Arabidopsis full-length cDNA library was constructed essentially
as reported previously (Seki et al., 1998, 2002). This clone is in a
modified pBluescript vector.
Please visit our web site (http://pfweb.gsc.riken.jp and
http://large.gsc.riken.jp) for further details.
reversed clone.

FEATURES
Location/Qualifiers
1..415

/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAFL07-95-D11"
/dev_stage="rosette plants"
/lab_host="DH10B"
/clone.lib="RAFL7"
/note="Site 1: BamHI; site 2: SalI; subjected to
cold-treated (1, 2, 5, 10, 24 hr)"

ORIGIN

Query Match 10.3%; Score 376.4; DB 3; Length 415;
Best Local Similarity 99.5%; Pred. No. 1e-85;
Matches 388; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

3224 ATAGGTTTGGATGGCGAATGCTGAAGATTGGACTGACAGACAGCGTGAACCTGCCA 3283
389 AGAGGTTTGGATGGCGAATGCTGAAGATTGGACTGACAGACAGCGTGAACCTGCCA 330
3284 GCTTGGGTTTATGATTATATACACTGTGAACTATCTGTGACAGTGTGACAGTCTC 3343
329 GCTTGGGTTTATGATTATATACACTGTGAACTATCTGTGACAGTGTGACAGTCTC 270
3344 AGCAGATGAAACCCGTGCTGTGTGGAAGCAACTGTGAGAGTCTGCTGTCTATCTGA 3403
269 AGCAGATGAAACCCGTGCTGTGTGGAAGCAACTGTGAGAGTCTGCTGTCTATCTGA 210
3404 TTTGGTTTCATCCAGAAACAATGCTACTGATGTGAGAACTTACACAACAAGATGAGAGT 3463
209 TTTGGTTTCATCCAGAAACAATGCTACTGATGTGAGAACTTACACAACAAGATGAGAGT 150
3464 TTTCTGTCGAAGTCAGGGTGGAAATCACTGAAGCTCTGTTCTTGATCATATATATAC 3523
149 TTTCTGTCGAAGTCAGGGTGGAAATCACTGAAGCTCTGTTCTTGATCATATATATAC 90
3524 TCATATGATGATGTCTGAGCTTGGAGATTCTCTTGTCTGTAATTTCTCTCTAAG 3583
89 TCATATGATGATGTCTGAGCTTGGAGATTCTCTTGT- TGTAATTTCTCTCTAAG 31
3584 TTAGGTTTATTAATGAACAAAAAATT 3613
30 TTAGGTTTATTAATGAACAAAAAATT 1

RESULT 7

B2437564/c 785 bp DNA linear GSS 13-DEC-2002
LOCUS BONRN72TR BO.1.6_2_KB_tot Brassica oleracea genomic clone BONRN72,
DEFINITION genomic survey sequence.

ACCESSION B2437564
VERSION B2437564.1 GI:26691135
KEYWORDS GSS.

SOURCE Brassica oleracea

ORGANISM Brassica oleracea

REFERENCE 1 (bases 1 to 785)
Ayala,M., Haas,B.J., Kumar,N., Wu,H., Xiao,Y., Van Aken,S.,
Utecherack,T.R., Wortman,J.R., White,O.R. and Town,C.D.
Whole genome shotgun sequencing of Brassica oleracea and its
application to gene discovery and annotation in Arabidopsis
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; euroside II; Brassicales; Brassicaceae; Brassica.

AUTHORS

TITLE JOURNAL
COMMENT Other GSSs: BONRN72TF
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TR

Class: sheared ends.
Location/Qualifiers
1..785
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="T01000D3"
/db_xref="taxon:3712"
/clone="BONRN72"
/clone_1db="BO 1.6 2 KB tot"
/note="Vector: pHOSt1; Site 1: BstXI; 1.6-2 kb sheared
total DNA inserted into pHOSt1 using BstXI linkers"

ORIGIN

Query Match 9.9%; Score 363.6; DB 9; Length 785;
Best Local Similarity 73.3%; Pred. No. 2.6e-82;
Matches 582; Conservative 0; Mismatches 109; Indels 103; Gaps 5;

1061 CTAGTGAAGGTTCTCGGGGCTCTCTGTATATTCAGAAAGGTGGTGAAGTGAAGTACTT 1120
Db CAAATGAAGGTTCTCGGGGCTCTCTGTATATTCAGAAAGGTGGTGAAGTGAAGTACTT 640
1121 CTTCGGTGGTGAAGGCTCTCTGTATATTCAGAAAGGTGGTGAAGTGAAGTACTT 1180
Db CTTCGGTGGTGAAGGCTCTCTGTATATTCAGAAAGGTGGTGAAGTGAAGTACTT 580
1181 GTTTTGAAGTGAAGGCTCTCTGTATATTCAGAAAGGTGGTGAAGTGAAGTACTT 1240
Db GTTTTGAAGTGAAGGCTCTCTGTATATTCAGAAAGGTGGTGAAGTGAAGTACTT 520
1241 CCGATTTTATTAATGCTGATATGAGTTGTTGAGAAAGTCTTGAAGCTTTACAGTACTT 1300
Db CCTATTTTATTAATGCTGATATGAGTTGTTGAGAAAGTCTTGAAGCTTTACAGTACTT 471
1301 TGACTGCTTGTGTAATTTGACGAGCGTGGCTTATTAAGAACTTTGATTGATCT 1360
Db 470 ----- 471
1361 TTGTTATTAAGTCTGTGTGAGAGAGAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1420
Db 470 -----TGCAGTTC 427
1421 CAAATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1479
Db CAGATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 367
1480 ACCCTGCTGATGATTAATGCTGCGAAAGAACTAAATGTTAAGCGGTGCGGAATAT 1539
Db ACCCTGCTGCGGATGAT-----AAACGACAGATGTTAAGCGGTGCGGAATAT 316
1540 TTGTGTGCTGTGTGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1599
Db CTTGTGTGCTGTGTGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 256
1600 GTTATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1659
Db ATTATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 196
1660 CTTTATTAATTTCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1714
Db TTTTATTAATTTCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 136
1715 TTGTGTGCTGTGTGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1774
Db TTGTGTGCTGTGTGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 76
1775 AAGTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1833
Db AAGTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 16
1834 CTTTATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1847
Db 15 CTTTATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2

RESULT 8
BP620404/c 434 bp mRNA linear EST 26-JUN-2004
LOCUS BP620404 RAFL16 Arabidopsis thaliana cDNA clone RAFL16-37-C18 3',
DEFINITION mRNA sequence.
ACCESSION BP620404
VERSION BP620404.1 GI:49271586
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Bukarota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 434)
Seki, M., Narusaka, M., Kamiya, A., Ishida, Y., Satou, M., Sakurai, T.,
Nakajima, M., Enju, A., Akiyama, K., Ono, Y., Muramatsu, M.,
Hayashizaki, Y., Kawai, J., Carninci, P., Itoh, M., Ishii, Y.,
Arai, K., Shibata, K., Shinagawa, A., and Shinozaki, K.
Functional annotation of a full-length Arabidopsis cDNA collection
Science 296 (5565), 141-145 (2002)
11910074
PUBMED
JOURNAL
TITLE *
COMMENT Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: meeki@rc.riken.go.jp
reversed clone, please visit our web site
(http://pfweb.gsc.riken.go.jp/) for further details.
Location/Qualifiers

FEATURES
source
1..434
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAFL16-37-C18"
/lab_host="DH10B"
/clone_1db="RAFL16"
/note="Site_1: BamHI; Site_2: SalI; dark-grown"

ORIGIN

Query Match 9.8%; Score 359.4; DB 3; Length 434;
Best Local Similarity 98.0%; Pred. No. 2.6e-81;
Matches 385; Conservative 0; Mismatches 6; Indels 2; Gaps 2;

3224 ATAGGTTTGAAGGGGGAATGCTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 3283
Db AGAGTTTGAAGGGGGAATGCTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 343
3284 GCTTGGTGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 3343
Db GCTTGGTGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 283
342 GCTTGGTGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 283
344 AGCAGATGAAGCCCGTCTCTGTGGAAGCACTCTGAGAGAGTCTGTTGATATGA 3403
Db AGCAGATGAAGCCCGTCTCTGTGGAAGCACTCTGAGAGAGTCTGTTGATATGA 223
282 AGCAGATGAAGCCCGTCTCTGTGGAAGCACTCTGAGAGAGTCTGTTGATATGA 223
3404 TTGTGCTTCAACCAAGAAACAACTGATGTCAGAACTTACACAAAGATTAAGCAAGT 3463
Db TTTTGTGCTTCAACCAAGAAACAACTGATGTCAGAACTTACACAAAGATTAAGCAAGT 164
222 TTTTGTGCTTCAACCAAGAAACAACTGATGTCAGAACTTACACAAAGATTAAGCAAGT 164
3464 TTTCGTGCTCAAGTCAAGGTGAAGAAATCA-CTGAAGGCTCTGTTCTGCAATTAATATA 3522
Db TTTCGTGCTCAAGTCAAGGTGAAGAAATCACTGAAGGCTCTGTTCTGCAATTAATATA 104
163 TTTCGTGCTCAAGTCAAGGTGAAGAAATCACTGAAGGCTCTGTTCTGCAATTAATATA 104
3523 CTCATATGAGCATGCTGAGCTTGGAGATTCCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 3582
Db CTCATATGAGCATGCTGAGCTTGGAGATTCCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 44
3583 GTTAGCTTTTAAATGAACAAATTAATTA 3615
Db 43 GTTAGCTTTTAAATGAACAAATTAATTA 11

RESULT 9
CX195205
LOCUS 13-B019365-021-009-J03-T7R ADIS-MP1Z 021 Brassica napus EST 28-DEC-2004
DEFINITION MP1ZP102200390, mRNA sequence.
ACCESSION CX195205
VERSION CX195205.1 GI:56842629
KEYWORDS EST.
SOURCE Brassica napus (rape)
ORGANISM Brassica napus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1 (bases 1 to 624)
AUTHORS Jakob, M., Lehmann, D. and Weishaar, B.
TITLE direct submission to GenBank (ADIS-MP1Z 021)
JOURNAL Unpublished (2004)
COMMENT Contact: Bernd Weishaar
Chair of Genome Research
Bielefeld University, Institute for Genome Research
Universitaetstrasse 25, D-33594 Bielefeld, Germany
Tel: +49-521-106 6873
Fax: +49-521-106 6423
Email: bernd.weishaar@uni-bielefeld.de
Insert Length: 624 Std Error: 0.00
Plate: 9 row: 3 column: 3
Seq primer: T7R CTATACGACTACTATAGGA.
Location/Qualifiers
1. 624
/organism="Brassica napus"
/mol_type="mRNA"
/cultivar="Expres 617"
/db_xref="GABI:1111775"
/db_xref="taxon:3708"
/clone="MP1ZP102200390"
/tissue_type="flowers and flower buds"
/dev_stage="young flowers and flower buds"
/clone_1lb="ADIS-MP1Z 021"
/note="Vector: pSPORT1, Site_1: SalI; Site_2: NotI; cDNA library from Brassica napus, strain Expres 617; RNA was from young flowers and flower buds of two greenhouse plants; library was made at the Max-Planck-Institute for Plant Breeding Research, Cologne, Germany; cloning sites SalI-NotI, primer sites and orientation:
T7-Sali-CCAGCGGTCGCG-5'utr-cDNA-polyA-CC-NotI-SP6, made by Marc Jakob 09/2000; PI: Bernd Weishaar. Sequence submission managed by R2PD/GABI-Primary database:
http://gabi.rzpd.de"

ORIGIN
Query Match 9.8%; Score 359.2; DB 8; Length 624;
Best Local Similarity 76.9%; Pred. No. 3.3e-81;
Matches 518; Conservative 0; Mismatches 93; Indels 63; Gaps 4;
Query 2035 AGGAATCCAGTATGTGAGTTGTTTGAAGAATTCAAAATCGATGACATGATGAT 2094
Db 2 AGGAATCCAGTATGTGAGTTGTTTGAAGAATTCAAAATCGATGACATGACAGC 61
Query 2095 CTCCCTGACTATGCAAAATGTGGAAACCTGTGGAGGAGGTTGCTTCTTAGGTTG 2154
Db 62 CTCCCTGAGCTCTGCAAAATTAATGGAGACCTGTGGAGGAGGTTGCTTCTTAGGTTG 121
Query 2155 AGAGACACCAAGATTAATACTCGGGGACTACTATGATGATGATGATGATGATG 2214
Db 122 AGAGATACCAAGATTAATACTCGGGGATTAATGCGATGATGATGATGATGATG 181
Query 2215 AGTACTGTGAAAGAGTGTGAGTGTGAGGTTCTCTTTAGCTGTGCTGCAACTATG 2274
Db 182 AGTACTGTGAAAGAGTGTGAGTGTGAGGTTCTCTTTAGCTGTGCTGCAACTATG 241
Query 2275 GCAAGGATTGAGCCGAGCATGTGAAGCTAGTGTCTATGACAGGCATGCAAGAAAGTTT 2334

Db 242 GCAAGGATTGAGCCGAGCATGTGAAGCTAGTGTCTATGACAGGCATGCAAGAAAGTTT 301
Query 2335 CCTTCCGCTATACAGATAGAAACTCGGCTGAACCCAGAGTGTGCAAGACAGCTTT 2394
Db 302 CCTTCCGCTATATCA-----GATGGCAAGAGACGGTGTCT 337
Query 2395 AGTGTAGATCTGTGTGTAACAATGATAGCCGTGA---TGTGAGCCGTGTGTCTTAAT 2451
Db 338 ACTGTAGATCATGTGTGTAACAATGATAGCATGATGATGATGATGATGATGATGATG 397
Query 2452 GCAGAGCTGTGAACCTCTGAAAACCTTTGAAACTAATGATTAATGCAATTCGAGCTGG 2511
Db 398 GCAGCAACTGAAGAGAGATCCGAATCTTTGATA----- 431
Query 2512 GTCTCAGAGATGACGTTGATGAAGAACTAGTGAATGTCCTGCTGATGATGATGAAG 2571
Db 432 ----CAAGTCTACGCTGATGATGATGATGATGATGATGATGATGATGATGATG 487
Query 2572 GAGGCAAGTGTGAAGATCTAGCTGTGTGTGCAATGTGACTGATTTCACTGTTCAGC 2631
Db 488 GAGGCAAGTGTGAAGATCTCTGTGCGGTGTGTGTGGAATAATTTGCTGCGCAGC 547
Query 2632 CAGAATATTTTCTTAAA-----GAGCTCATCTTTTCAAGCAGATATGCTTCT 2685
Db 548 CTGAGGATATGTCTATTAAGGCGGAGCATGATGATGATGATGATGATGATGATG 607
Query 2686 TCTATGGAATCTGA 2699
Db 608 TCTAGGATCTGA 621

RESULT 10
CD573714
LOCUS 897 bp mRNA linear EST 12-JUN-2003
DEFINITION UCRPT01_01_P12.T3 Poncirus trifoliata CTV-challenged cDNA library -
Poncirus trifoliata cDNA clone UCRPT01_01_P12, mRNA sequence.
ACCESSION CD573714
VERSION CD573714.1 GI:31669616
KEYWORDS EST.
SOURCE Poncirus trifoliata
ORGANISM Poncirus trifoliata
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Sapindales; Rutaceae; Poncirus.
REFERENCE 1 (bases 1 to 897)
AUTHORS Roose, M.L., Ye, X., Federici, C.F., Close, T.J., Fenton, R.D., Wananaker, S., Choi, Y. and Kingan, T.
TITLE Development of EST Resources and New Genetic Markers for California Citrus - Poncirus trifoliata CTV-challenged phloem - UCR Unpublished (2003)
JOURNAL Contact: Mikeal Roose
Department of Botany & Plant Sciences, University of California
Riverside, CA, 92521-0124, USA
Tel: 909/8744137
Fax: 909/8744137
Email: mikeal.roose@ucr.edu
Seq primer: T3.
FEATURES
SOURCE
1. 897
/organism="Poncirus trifoliata"
/mol_type="mRNA"
/cultivar="Pomeroxy OP"
/db_xref="taxon:37690"
/clone="UCRPT01_01_P12"
/tissue_type="phloem"
/dev_stage="10 - 30 cm shoots"
/lab_host="E. coli TUC121"
/clone_1lb="Poncirus trifoliata CTV-challenged cDNA library - UCR"
/note="Vector: Lambda Uni-ZAP XR, excised phagemid; Site_1: EcoRI; Site_2: XhoI; Plants were grown in the greenhouse at University of California, Riverside. the

scion was a open-pollinated (very probably selfed) seedling of *Poncirus trifoliata* cv Pomeroy that was selected as homozygous for the Ctv resistance gene. The rootstock was sweet orange infected with citrus tristeza virus (CTV) isolate 7514 over 1 year before sampling (CTV infects sweet orange, but not genotypes carrying the Ctv resistance gene. Shoots 10-30 cm long were harvested in October 2000, and the green phloem (bark) was removed and frozen quickly in dry ice. Total RNA was extracted using Trizol reagent (Gibco). Poly(A) RNA was purified, a cDNA library was made, and 0.5 million primary lambda cDNA clones were in vivo excised to give a population of phagescript SK(-) phagemids. All steps to this point were performed in the M. Rouse lab at the University of California, Riverside by X. Ye. Phagemids were plated, plasmid DNA purified, cDNA clones archived, and DNA sequences determined bi-directionally using an ABI3730 at the University of California Riverside Genome Institute, Core Instrumentation Facility, (Choi, Kinsan). Chromatogram files were transmitted to UC Riverside (by Choi), then processed at UC Riverside (by Manamaker) using the Harvest pipeline (<http://harvest.ucr.edu>) to remove contaminants, and to trim to a high quality region. Sequences that retained a phred 17 region of at least 100 bases were deposited to GenBank."

ORIGIN

Query Match 9.2%; Score 336.4; DB 6; Length 897;

Best Local Similarity 63.9%; Pred. No. 2.9e-75;

Matches 574; Conservative 0; Mismatches 241; Indels 83; Gaps 1;

```
QY 1391 GCAAGTACCTTGGACCGGATTTACGACAAATTGATGAGCTTTGGAGAGATCACT 1450
DB 80 GCCAGTGTCTCGCTCCAGATTGGACGACAGATTGACGAGCTTGGAGAGATCACT 139
QY 1451 CCGGCTTATGCTTGGAGCTACTTGGCTTACCGCTTGTGATGATTCAGCTGCAAGAA 1510
DB 140 CCAGCGCTGTCTGGAACCTTTAGGCTTACCCCTTACTGTGATGATTCAGCAAGAGA 199
QY 1511 CTAATAGTGTAAAGCGGTGGGAATTTTGTGCTGTGTGAGAGAGGTGGAGCATCA 1570
DB 200 GAAAGAGACTTCATGATATCTCAACTACTATGAGCTGTGTGAGAGAGGAGCACT 259
QY 1571 GCTCTTGTGGGGGTTTGACCCGAGAGATTATGATGAGCGGTTTTCAGATGCA 1630
DB 260 GCATTGTGCGGGGATTACACGTAATCTTCAATGATGAGGCTTTCTACATGACA 319
QY 1631 GCTGCTGACAGATATACAGTTTAAATTTCTTTAGCATGATATACT 1690
DB 320 TCAGCTGAGC----- 329
QY 1691 TTAGGTTTCAATTTATGATATGTTGTGTAGGTTGATCTTTGTAGTACCCCAA 1750
DB 330 -----AGTTTAACTATTTCTCCACACCAA 356
QY 1751 GCAATATTCACAGAGATCATTTGAGTTTGAAGTTGACATTGCTTGTGCTCAAG 1810
DB 357 ATAGCATCCGAGCTGAACCTTTGAGCTTATGAGTGGCATTTGCACTTTTCCCAAG 416
QY 1811 CTTTATTTGTAAGAACCACTTTTACAGATGCTGATTAAGCAATTCAGCAACTTC 1870
DB 417 CTTTGTGGTGAAGAACCTCATCTCATTTGAGATGCTGATTAACATGTTCAAGCACTTC 476
QY 1871 AGCAGGCTAAGTATGCTATGAGATTCCTGCGATTTGTATGATTAACAGGAATATT 1930
DB 477 AGCAGAAATTAAGTATCCAGCTTTAGCGGACCTTGTCTATATTAATTCCTTGAAGAAAC 536
QY 1931 GGGAGATAGACTTGGCTTGAAGAGGGGACTGTGCACTGCTTATAGGCAAGTTGATG 1990
DB 537 ATGAGATGAGATTTGCATTTGGAAGGGGCTTATGTCTCTGCTTGTAGATAGCTTGATG 596
QY 1991 AATGCCGTAATGTGTTGGGCTTGAACAGTGAAGATTCAATATTAAGAAATCCAGCTATTG 2050
```

```
DB 597 AGTCCCTTATGATGTTGGGCTTATGATGATTAAGTCCCATACAGAAATCCAGCATTTG 656
QY 2051 TGGAGTTGTTTGGAGAAATTCAAATGCTGATGACATGATATTCCTCCGAGCTATGCA 2110
DB 657 TAGTTTGTGTTGGAAACTCAAGAGAACTATGACATGATATTCCTCGTCTCTGTA 716
QY 2111 AATGTTGGAACCTGCTTGGAGAGGGGTTGCTTCTTCTAGTTTCAGAGACCAAGATATA 2170
DB 717 AGCTGTAGAGACATGCTGCGAGAGAGTTGTTTCCCTTACATTCAGGACACCAAGGATA 776
QY 2171 AAAATTAACTCGGGGACTTACTATGATATCTTATGTTTGAAGTTTGGAAAAGAG 2230
DB 777 TACGTTTCAAACTGGTGATTAATGATATGATATCTTCTTCCATTTATTTGAAAAGAC 836
QY 2231 TGGAGTAGTTTCAAGGGTTTCTTTAGCTGCTGTCGCAATATGCAAGAGATTGAGAGC 2288
DB 837 TGGAGGAAATGTGATGTTCCCTTACTGACAGCAGCTGCCATATGTTAGGCTTGAGC 894
```

RESULT 11

DU050213/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

FEATURES

source

1..871

/mol_type="genomic DNA"

/cultivar="Heinz 1706"

/db_xref="taxon:4081"

/clone="LE_HBa0013821"

/lab_host="E. coli"

/clone_lib="Tomato HindIII BAC library"

/note="Vector: pBelobAC11, Site_1: HindIII"

ORIGIN

Query Match 9.1%; Score 334; DB 10; Length 871;

Best Local Similarity 67.4%; Pred. No. 1.2e-74;

Matches 503; Conservative 0; Mismatches 235; Indels 8; Gaps 2;

```
QY 1374 TTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1433
DB 791 TTTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 732
QY 1434 CTTTGAAGAGATCACTCCGCTTATGCTTGAAGCTTACTTGGCTTACCGCTTGATG 1493
```

Db		731	CATTGGAAGATTAATCAACGGCTACGTACTTGAACCTTCAGCTTTCCCTTGCGTAGTC	672
Oy		1494	AATTACGCTGCAGAAAGACTAAATGEGTTTAAAGCGGTGTCGGAATAATTTGTGCTCTGTTG	1553
Db		671	AATACCGAATGAAAAGAGTAGAGGCTCTTCAAAGGTGTGCGCAACAATTTTGTGCGCTGTTG	612
Oy		1554	GAGGAGGTGTGAGCATCACTCTTGTGTGGGGGTTTACCCTGTGAGAAATTATGAAATGAG	1613
Db		611	GAGGAGGAGGAGCAGACTCAATTTTCAGGGGGGGTTCACAGAAAGATTTCATGAATGAG	552
Oy		1614	CGTTTTTACGATGACAGCTGCTGAGCAGG-----TATACAGTTTAGAATACCTTTTATTA	1668
Db		551	CTTCTCTACGATGACAGCTGTGTGAGCAGGTTGGCTTACATTCAATTATTTTGGTT	492
Oy		1669	ATTCTTTAGCATGATATTAACCTTTAGGTTTCTCAATTTAATGATGTTGTGTAGAGTT	1728
Db		491	TCTTTTGTAACTGTTCTTTTCACTCCTTCTCTTCAATTTTCCTTTCTTTTAAAGGTG	432
Oy		1729	GATCTTTTGTGAGCTAACCCCAAGCAATTTCCAGAGAGTCAATTGAAAGTTTACGAATT	1788
Db		431	GACCTCTTCGTGCAACGCCCAAGTAACATTCCTGCAAGAAAGCTTGAAGTTTACGGGGTG	372
Oy		1789	GCACTTGCTCTGTGGGCTCAAGCTTTTATTTGTTAAGAACCAACCTTTACAGAGTCT	1848
Db		371	GCGCTTGACATTTGTCTCAAGACTTTTGTGGAAAAAACTCATCTCATTCAGATATCT	312
Oy		1849	GATAAGCAATTCACAGCACTTCAGCAGGCTAAAGTAAATGCTATGAGATTTCTGCGATG	1908
Db		311	GATAACCTTTTTCAGCACTTCAGAGACCAAAAGTTACAGCTTACGGCAGTTCTGTGTCT	252
Oy		1909	TTGTATGTATCACGGAATATTTGGAGATPAGACTTCGCTCTAAGAAAGGGAACCTGTGCA	1968
Db		251	GTATPACACTGTTAAGAAAAACGTTGAATPAGACTTCGCTTTGAGAGGGGCTTTGTCCA	192
Oy		1969	CTGCTTATAGCCAAAGTTGATGAATGCCGTATGCGTTGGGCTTAAAGATGAGGATTTCA	2028
Db		191	CTGCTTTGTGAGAGGTGATGAAATGTGCTCTTGTGTGGGCTTAAGACATGAGGACATCC	132
Oy		2029	CAATATAGAAATCCAGACTATTGTGGAGTTTGTTTGGAGAAATTCGAATCGTATGAC---	2085
Db		131	CTTATAGAGATCCAGCTATAGTACTTTTGTGCAAGACACTCAAGAGACGACACGGA	72
Oy		2086	AATGATGATCTCCCTGCACTATGCAA	2111
Db		71	AATGATCTGCTCCCTGCACTATGTA	46
RESULT 12				
B0834167		534 bp	mRNA	linear EST 14-MAR-2003
LOCUS		ALSTO128	Arabidopsis lyrata Inflowerence PCMV-POR Library	
DEFINITION		Arabidopsis lyrata cDNA clone PIMB1-D03 3', mRNA sequence.		
ACCESSION		B0834167		
VERSION		B0834167.1	GI:28951482	
KEYWORDS		EST.		
SOURCE		Arabidopsis lyrata		
ORGANISM		Arabidopsis lyrata		
		Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
		Spermatophytes; Magnoliophyta; eudicotyledons; core eudicotyledons;		
		rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.		
REFERENCE		1 (bases 1 to 534)		
AUTHORS		Barrier,M., Bustamante,C.D., Yu,J. and Purugganan,M.D.		
TITLE		Selection on rapidly evolving proteins in the Arabidopsis genome		
JOURNAL		Genetics 163 (2), 723-733 (2003)		
PUBMED		12618409		
COMMENT		Contact: Barrier M		
		Department of Genetics		
		North Carolina State University		
		3513 Gardner Hall, Box 7614, Raleigh, NC 27695, USA		
		Tel: 919 515 1761		
		Fax: 919 515 1695		
		Email: mdbarrle@unity.ncsu.edu		

FEATURES					
source	Plate: 1 row: H column: 5 Seq primer: T3. Location/Qualifiers 1..534 /organism="Arabidopsis lyrata" /mol_type="mRNA" /culivar="Karhumaki" /db_xref="taxon:59689" /clone="PIWB1-D03" /issue_type="Inflorescence" /clone_idb="Arabidopsis lyrata Inflorescence pCMV-PCR library" /note=Vector: pCMV-PCR (Stratagene); Created using PCR Library Construction Kit (Stratagene)"				
ORIGIN					
Query Match	8.6%;	Score 315.8;	DB 5;	Length 534;	
Best Local Similarity	76.9%;	Pred. No. 5.4e-70;			
Matches 442;	Conservative 0;	Mismatches 42;	Indels 91;	Gaps 1;	
Dy	2971	TTCTTCCCTTGANTTTTTTCCTTTTCGATTTAGGGTCGTCAAGCTGCATTCAGAGC	3030		
Db	51	TTCTTATGGAATCTGACATTGCTACCATAGGGTCAGTAGACTGATTCAGAAGC	110		
Dy	3031	ACTTCCAGAAATGAGTCTAGACTGAGAGAAATATATGTATCCAAGTGGCAAGATTA	3090		
Db	111	ACTTCCAGATGAGTCTAGACTGAGAGAGATATGTATCCAAGTGGCAAGATCAA	170		
Dy	3091	GTCCTCGGCTTTGGGGCTGATCACCGCATAGAAATGTATCCAAGGTGAGGGAAATAAT	3150		
Db	171	GTCCTAGGCTTTGGGATATATCATCTGATAGAAATCTATCAG-----	214		
Dy	3151	CTACAATTCATCATATGTGTGAAAACTGTGACATGATTAATAGTGTGCTGCTGTTT	3210		
Db	215	-----	214		
Dy	3211	GATTCTGTATTTATTAAGCTTTTGAGTGGCGAATGCTGAAGATTTGCATGACAGACAG	3270		
Db	215	-----AGTTTTGGATGGGGGAATGCTGAAGATTTGGCATGACAGAGCTG	259		
Dy	3271	CTGAATCTGGGAGCTTGGGTTGGTTTATGATTAATACATGTTGAACATACTGTTGACA	3330		
Db	260	TGGAATCTGGGAGCTTGGGTTGATTAATTAATACATGTTGAACATACTGTTGACA	319		
Dy	3331	GTCGACAGTCTGACAGATGAGAACCGGCTCTGTGTGAAGCAACTGAGAGAGCTG	3390		
Db	320	GTCGACAGTCTGACAGATGAGAACCGGCTCTGTGTGAAGCAACTGAGAGAGCTG	379		
Dy	3391	CTGTCTATCTGATTTGGTTCAATCCAGAAAACAATGCTAGTGTGAGAACTTACAA	3450		
Db	380	CTGTCTATCTGATTTGGTTCAATCCAGAAAACAATGCTAGTGTGAGAACTTACAA	439		
Dy	3451	CAAGATCGAAGTTTTCTGGTCCAAGTCAGAGGTGGAAAATCACTGAAGGCTCTGTTCTG	3510		
Db	440	CAAGATCGAAGTCTGTGTCCAAGTCAGAGGTGGAAAATCACTGAAGGCTCTGTTCTG	499		
Dy	3511	CATCATATATTAATCATATGATGACATGCTGAGCT	3545		
Db	500	CATCATATATGCTCTTATATAGCATACCAGACT	534		
RESULT_13					
AVB30764	320 bp	mRNA	linear	EST 01-APR-2002	
LOCUS	AVB30764	RAFL9	Arabidopsis thaliana cDNA clone RAFL09-76-G11 5'		
DEFINITION	mRNA sequence.				
ACCESSION	AVB30764				
VERSION	AVB30764.1	GI:19872824			
KEYWORDS	EST.				
SOURCE	Arabidopsis thaliana (thale cress)				
ORGANISM	Arabidopsis thaliana				
	Brassicaceae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;				

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

roside; eurosida II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 320)
Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M.,
Oono,Y., Sakurai,T., Carninci,P., Kawai,J., Itoh,M., Ishii,Y.,
Arakawa,T., Shibata,K., Shingawa,A., Muramatsu,M., Hayashizaki,Y.
and Shinozaki,K.
Large scale analysis of Arabidopsis full-length cDNA (2002b)
Unpublished (2002)
Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: msekic@r.riken.go.jp
An Arabidopsis full-length cDNA library was constructed essentially
as reported previously (Seki et al., 1998). cDNA cleaved with BamHI
and XhoI was ligated to modified lambda FLc-1 vector (Carninci et
al., submitted for publication) digested with BamHI and SalI. This
clone is in a modified Bluescript vector. Please visit our web
site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further
details.

FEATURES
source
1..320
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAF109-76-G11"
/dev_stage="plants at various developmental stages from
germination to mature seeds"
/lab_host="DH10B"
/clone_1lb="RAF19"
/note="Site 1: BamHI; Site 2: SalI; subjected to
dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24
hr) treatments"

ORIGIN

Query Match 8.5%; Score 311.8; DB 1; Length 320;
Best Local Similarity 98.4%; Pred. No. 5e-69;
Matches 313; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 368 GATTAACTTACTACTCAAAATCCATAACCTTACACCAACAGCTC 427
DB 2 GATTAACTTACTACTCAAAATCCATAACCTTACACCAACAGCTC 61
QY 428 TTCAATATGTAACAGAAAGTTTGTAGTAGCTTAAAGACACTCCATGAG 487
DB 62 TTCAATATGTAACAGAAAGTTTGTAGTAGCTTAAAGACACTCCATGAG 121
QY 488 CTCTGAGTCAGCTGGCATTGGTCTCTCCCAATTCATTATGCGATTACACCGCGGA 547
DB 122 CTCTGAGTCAGCTGGCATTGGTCTCTCCCAATTCATTATGCGATTACACCGCGGA 181
QY 548 CGAAGAGCTCCGACGTAGCAACACCTTACAACTATGCTCCGCGCAAGATGG 607
DB 182 CGAAGAGCTCCGACGTAGCAACACCTTACAACTATGCTCCGCGCAAGATGG 241
QY 608 CCGACCGCTCTCTCCGACTTCAATTCACCTCCGATTCCTCTCCCTCTTCCGCA 667
DB 242 CCGACCGCTCTCTCCGACTTCAATTCACCTCCGATTCCTCTCCCTCTTCCGCA 301
QY 668 CCGCACCACCAACCGCCA 685
DB 302 CCGCACCACCAACCGCCA 319

RESULT 14
BZ765587 320 bp DNA linear GSS 13-MAR-2003
LOCUS BZ765587
DEFINITION SALK_131837.28.70.x Arabidopsis thaliana TDNA insertion lines
Arabidopsis thaliana genomic clone SALK_131837.28.70.x, genomic
survey sequence.
ACCESSION BZ765587

VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BZ765587.1 GI:28938140
GSS.
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosida II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 320)
Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R.,
Gadrinab,C., Jeske,A., Karnes,M., Kim,C.U., Parker,H., Prednis,L.,
Shim,P., Zimmerman,J. and Ecker,J.R.
A Sequence-Indexed Library of Insertion Mutations in the
Arabidopsis Genome
Unpublished (2001)
Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGnAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu
This is single pass sequence recovered from the left border of
TDNA. This sequence lies within 300 bases of the 5' end of
AT5G42480.
Class: TDNA tagged.

FEATURES
source
1..320
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/ecotype="Col-0"
/db_xref="taxon:3702"
/clone="SALK_131837.28.70.x"
/note="PCR was performed on Arabidopsis thaliana lines
each of which contains one or more TDNA insertion
elements. The resultant fragment for each line was
directly sequenced to determine the genomic sequence at
the site of insertion. Details of the protocols used can
be found at http://signal.salk.edu/tDNA_protocols.html"

ORIGIN

Query Match 8.4%; Score 307.2; DB 9; Length 320;
Best Local Similarity 97.5%; Pred. No. 7.6e-68;
Matches 312; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 343 TTTTCAGTACCTGCGATGTTTGTGATTTAACTTACTACTCAAAATCAAAATTCAT 402
DB 1 TTTTCAGTACCTGCGATGTTTGTGATTTAACTTACTACTCAAAATCAAAATTCAT 60
QY 403 AAACCTTAGACGACCAACAGTCTCTTCAATATGTAACAGAAACAAAGTTTGTAGTA 462
DB 61 AAACCTTAGACGACCAACAGTCTCTTCAATATGTAACAGAAACAAAGTTTGTAGTA 120
QY 463 GCGTAAATAACACCTCCAGTAGAAGCTCTGAGTCAGTGGCATTGGTCTCCCATTC 522
DB 121 GCGTAAATAACACCTCCAGTAGAAGCTCTGAGTCAGTGGCATTGGTCTCCCATTC 180
QY 523 CAATTAATGCCGATTACCAACCGGCGACGACCAAAAGCTCGAGTAGCCACAACCTCTACA 582
DB 181 CGATTATGCCGATTACCAACCGGCGACGACCAAAAGCTCGAGTAGCCACAACCTCTACA 240
QY 583 ACTATCTGCTCCGACGACCAAAATGGGCGACCGTCTCTCCGACTTCAATTTGACCTCC 642
DB 241 ACTATCTGCTCCGACGACCAAAATGGGCGACCGTCTCTCCGACTTCAATTTGACCTCC 300
QY 643 GATTCCCTCCCTCCCTCTT 662
DB 301 GATTCCCTCCCTCCCTCTT 320

RESULT 15
CG961431/c 954 bp DNA linear GSS 15-DEC-2003
LOCUS CG961431/c

DEFINITION	MBEKh28TR mch2 Medicago truncatula genomic clone 73F8, genomic survey sequence.
ACCESSION	CG961431
VERSION	CG961431.1 GI:39883077
KEYWORDS	GS.
SOURCE	Medicago truncatula (barrel medic)
ORGANISM	Medicago truncatula
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta, eudicotyledons, core eudicotyledons; rosids; eutrosids I; Fabales; Fabaceae; Papilionoideae; Trifoliales; Medicago.
AUTHORS	1 (bases 1 to 954)
TITLE	Town, C.D., Shetty, J., Koo, H. and Feldblyum, T.F.
JOURNAL	Sequencing of BAC ends from Medicago truncatula
COMMENT	Unpublished (2003)
	Other GSSs: MBEKh28TRFB
	Contact: Chris Town
	TIGR
	9712 Medical Center Drive, Rockville, MD 20850, USA.
	Tel: 301-838-3523
	Fax: 301-838-0208
	Email: cdtown@tigr.org
	Seq primer: CAGGAACAGCTATGACC
	Class: BAC ends.
FEATURES	
source	Location/Qualifiers
	1..954
	/organism="Medicago truncatula"
	/mol_type="genomic DNA"
	/culivar="genotype A17"
	/db_xref="taxon:3880"
	/clone="73F8"
	/clone_id="mch2"
	/note="Vector: pBelBAC11; Site_1: HindIII; Site_2: HindIII; Cook, D.R. and Kim, D.J., unpublished"
ORIGIN	
Query Match	8.3%; Score 305.6; DB 10; Length 954;
Best Local Similarity	65.8%; P: Mismatch 239; Indels 54; Gaps 6;
Matches 563; Conservative	0; Mismatches 239; Indels 54; Gaps 6;
1377	TGTAGAGGAAGAGCAAGTAGCCTTGACACCGGATTTACGTGCAAAATTGATGAGACTT 1436
856	TGTAGAGGAAGAGGCGCAAGCCTTAGCACCGGATTTCAAAACCAATTTATGAGACAC 797
1437	TGGAAGATCACTCCGCGTATGCTTGGAGCTTACCTTACCGCTTGATGATTT 1496
796	TTGAAGATTAACCTCAAGTTGTGATTTAGAACTTTAGCCTTCCTCTGATGATGAC 737
1497	ACGCTGGAAAGACTAAATGCTTTAAACCGGTGCGGATATTTTGTGCTGTGGAG 1556
736	ATCAGAGCGGAGGAAGAAAGTCTCCAAAGGTGTCGCAACATTTCTGGCAGATTGGAG 677
1557	GAGGTGAGCATGAGCTCTTGTGGGGGTTTGACCCCGGAGAAAGTTATGAAAGAGCGT 1616
676	GTGAGGAGCAGCAGCAATGCTGGAGTTTCAACGAGAGATTTCAATGATGAGCAT 617
1617	TTTACAGATGACAGCTGCTGAGCAGGATATACGTTTAGATA-----CCTTTTAT 1670
616	TCTGCGATGAAAGCTGCCGAACGGTTCATGAAATATATTTGTCATGTTTTTAT 557
1671	TTCTTAGCATATATACTTAGGTTTCTCATTTATATGAT----- 1713
556	TTAATTTTTCGCCCCCTTAGTTTGTATTTGAAGGTTTATACATTTATCTTTTCA 497
1714	-----GTGTGTGTAGTTGATTTTGTAGTATCCCAAG 1751
496	TTAATCTTTTATCTCAATCTTTTATTTTAAAGTGAATTTTGTAGCAACCTAG 437
1752	CAATATTCAGACAGATCATTTGAAGTTTACGAAGTTGACCTTCTGTGGGCTCAAG 1811
436	CAATATTCAGCTGAAGTTTGTAGGCTTATGGGTTGCACTTGATTTGTTGCAAGC 377
1812	TTTTATTGTGAGAGCAACCTTTTACAGATCTGATTAAGCAATTTCCAGCACTTCA 1871

Db	376	CTTTGTAGGTAAAAAGCCAACTCTTATCCAGATGCTGATTAATTATTCATCAACTTCA	317
Qy	1872	GCAAGCTAAGATATGCTATGAGATTCTGCAGTGTGTATGATAC-----ACGGAA	1925
Db	316	ACAACTAAGGTAAACAATATGAGGATGCTCCTCTGTTATCTCCATGAGAGATGA	257
Qy	1926	TAATTGGAGATAGACTTCGCGTCTAGAAAGGGACTCTGTGCACCTGTTATAGCAANGT	1985
Db	256	GAAGAAGAAGTGAATTTGTCATTAGAAAGGGGCTGTGTGCACCTGCTGTGGGAGCT	197
Qy	1986	TGATGAATGCCCGTATGTGCTTGGGCTTAGACAGTGA-GAATTCACAATATBGAATCCAG	2044
Db	196	TGATCAATGTGCATCACTAGTTGGGCTCTGCATAGACAGTCTCACTTATCGAAACCCAT	137
Qy	2045	CTATGTGGAGTGTGTTTGGAGAAATTCAAATGCTGATG-ACATGATGATCTCCCTCGA	2103
Db	136	CTATTATAAGCTTATATTATGAAAAACGAAAGGTGATGAAGTACAGTGAATCTTCCCTGA	77
Qy	2104	CTATGCAAAATGTGTGAAACCTGGTTGGCAGGGGTTCTTTTC-CTAGGTTGAGAGAC	2162
Db	76	CTCTGTAATTTGTTGGACATGGTGTGATGGAGGTGTTTTTCCCTAAGGTTTAGAGATAC	17
Qy	2163	CAAAAGATPAAAAATTT 2178	
Db	16	TAAGAGCAAACTTT 1	

Search completed: December 11, 2005, 10:07:53
Job time : 13954.2 secs

Qy	2163	CAAAGATAAAAAATT	2178
Db	16	TAAAGAGACAAACTTT	1

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 10, 2005, 08:09:29 ; Search time 2074.73 Seconds
(without alignments)
11779.592 Million cell updates/sec

Title: US-10-600-070B-3

Perfect score: 3667

Sequence: 1 tgcctgcatacgaagagaat.....ctacacatacgaagcctacaa 3667

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

N_Geneseq_21:*

- 1: geneeqn1980s:*
- 2: geneeqn1990s:*
- 3: geneeqn2000s:*
- 4: geneeqn2001as:*
- 5: geneeqn2001bs:*
- 6: geneeqn2002as:*
- 7: geneeqn2002bs:*
- 8: geneeqn2003as:*
- 9: geneeqn2003bs:*
- 10: geneeqn2003cs:*
- 11: geneeqn2003ds:*
- 12: geneeqn2004as:*
- 13: geneeqn2004bs:*
- 14: geneeqn2005as:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3667	100.0	3667	ADJ38130	Arabidops
2	3663.8	99.9	3667	ADJ38136	Arabidops
3	1968.4	53.7	2679	ADT14901	Plant cdn
4	1931	52.7	2637	ADJ38210	Plastid d
5	1724	47.0	2406	ADJ38129	Arabidops
6	1720.8	46.9	2406	ADJ38135	Arabidops
7	1717.6	46.8	2406	ADJ38208	Plastid d
8	402.2	11.0	561	ADJ38212	Plastid d
9	295.6	8.1	631	ADJ38264	Plastid d
10	208.8	5.7	660	ADJ38215	Plastid d
11	194	5.3	2283	ADJ38206	Plastid d
12	178.6	4.9	607	ADN48781	Cotton pr
13	175	4.8	439	ADP93143	Cotton ex
14	161	4.4	545	ADJ38255	Plastid d
15	159.4	4.3	552	ADN48855	Cotton pr
16	158	4.3	608	ADJ38217	Plastid d
17	145.6	4.0	527	ADJ38214	Plastid d
18	136.2	3.7	647	ADJ38265	Plastid d
19	134.6	3.7	652	ADJ38266	Plastid d

C	20	132.2	3.6	563	12	ADJ38258	Adj38258 plastid d
C	21	130.2	3.6	446	12	ADJ38231	Adj38231 plastid d
C	22	127.6	3.5	479	12	ADJ38230	Adj38230 plastid d
C	23	125	3.4	537	12	ADJ38223	Adj38223 plastid d
C	24	122.6	3.3	307	12	ADJ38218	Adj38218 plastid d
C	25	118.6	3.2	871	12	ADJ38232	Adj38232 plastid d
C	26	93.6	2.6	491	12	ADJ38254	Adj38254 plastid d
C	27	92.6	2.5	187	12	ADJ38216	Adj38216 plastid d
C	28	92	2.5	418	12	ADJ38224	Adj38224 plastid d
C	29	91.4	2.5	460	12	ADJ38267	Adj38267 plastid d
C	30	90.4	2.5	420	12	ADJ38256	Adj38256 plastid d
C	31	89.4	2.4	604	12	ADJ38227	Adj38227 plastid d
C	32	89.4	2.4	653	12	ADJ38228	Adj38228 plastid d
C	33	87.2	2.4	480	12	ADJ38225	Adj38225 plastid d
C	34	86.8	2.4	622	12	ADJ38226	Adj38226 plastid d
C	35	85.2	2.3	300	12	ADJ38260	Adj38260 plastid d
C	36	81.8	2.2	360	12	ADJ38259	Adj38259 plastid d
C	37	80.8	2.2	606	12	ADJ38257	Adj38257 plastid d
C	38	78	2.1	416	12	ADJ38219	Adj38219 plastid d
C	39	77.2	2.1	549	12	ADJ38261	Adj38261 plastid d
C	40	75.8	2.1	340	12	ADJ38233	Adj38233 plastid d
C	41	66.6	1.8	336	12	ADJ38222	Adj38222 plastid d
C	42	65.6	1.8	309	12	ADJ38221	Adj38221 plastid d
C	43	60.2	1.6	295	12	ADJ38213	Adj38213 plastid d
C	44	59.4	1.6	400	12	ADJ38263	Adj38263 plastid d
C	45	54	1.5	535	12	ADJ38229	Adj38229 plastid d

ALIGNMENTS

RESULT 1
ID ADJ38130 standard; DNA; 3667 BP.
XX ADJ38130;
AC
XX
DT 06-MAY-2004 (first entry)
XX
DE Arabidopsis thaliana Arc6-1 genomic DNA SegID2.
XX
XX prokaryotic type; plastid division; Ftn2; ARCE; ARCS; Fzo; plant cell;
XX agronomic; horticultural; crop plant; ornamental plant; woody plant;
XX herbicide target; ds.
XX
OS Arabidopsis thaliana.
XX
XX WO2004001003-A2.
XX
XX
XX
XX 31-DEC-2003.
XX
XX 20-JUN-2003; 2003WO-US019536.
XX
XX
XX 20-JUN-2002; 2002US-0390140P.
XX 09-AUG-2002; 2002US-040242P.
XX 20-JUN-2003; 2003US-0060070.
XX
XX (UNMS) UNIV MICHIGAN STATE.
XX
XX Ostrya Young KW, Vittha S, Kokscharova OA, Gao H;
XX WPI; 2004-082486/08.
XX P-PSDB; ADJ38202.
XX
XX
XX New isolated Ftn2, ARCS and/or Fzo-like nucleic acid sequences, useful
XX for further characterizing plastid division in plant cells, and in
XX varying agronomic and horticultural characteristics of economically
XX important plants.
XX
XX Claim 1; SEQ ID NO 2; 287pp; English.
XX
XX This invention relates to novel prokaryotic type or plastid division and
XX related genes and proteins. In particular, the invention relates to novel

CC Ftn2 (AR6), AR6 and Fzo-like genes and polypeptides. The methods and
CC compositions of the present invention are useful for further
CC characterizing the present invention in plant cells, in order to vary
CC agronomic and horticultural characteristics of economically important
CC plants, such as crop, ornamental and woody plants. They can also be used
CC as herbicide targets. The present sequence is that of a DNA sequence
CC which is related to the invention.

XX Sequence 3667 BP; 983 A; 670 C; 842 G; 1172 T; 0 U; 0 Other;

Query Match 100.0%; Score 3667; DB 12; Length 3667;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 3667; Conservative 0;

```
QY 1 TGTTCGTGATTAAAGAGAAATACAAATTAATAGCAATTTGCTTGATTTACAGAAATTTTG 60
DB 1 TGTTCGTGATTAAAGAGAAATACAAATTAATAGCAATTTGCTTGATTTACAGAAATTTTG 60
QY 61 CTTCGCTATAGGATTCATTGGCTCTGTTTGCTTTTACATTTACATGTCATATAGTTTCG 120
DB 61 CTTCGCTATAGGATTCATTGGCTCTGTTTGCTTTTACATTTACATGTCATATAGTTTCG 120
QY 121 AATTTTACACATTTGATTTGATTTAAGAAAGAGAGAAATGATGGGGTTTTTGCGG 180
DB 121 AATTTTACACATTTGATTTGATTTAAGAAAGAGAGAAATGATGGGGTTTTTGCGG 180
QY 181 TTTAACTTAAAGTACAGAAATTAAGTCATTTGCTTCTATATGCTTAA 240
DB 181 TTTAACTTAAAGTACAGAAATTAAGTCATTTGCTTCTATATGCTTAA 240
QY 241 AATGAGGCACTCCACAGGTTCTTAGTGATATGATTTATAGACATTTAACATCAT 300
DB 241 AATGAGGCACTCCACAGGTTCTTAGTGATATGATTTATAGACATTTAACATCAT 300
QY 301 AAAAGCGTGAGGCACTGTAACATCATAGTTGTTTTTATTTTTCAGTAGCTGTGAT 360
DB 301 AAAAGCGTGAGGCACTGTAACATCATAGTTGTTTTTATTTTTCAGTAGCTGTGAT 360
QY 361 GTTTTGTATTTAACTTAACTTAACTCAAAATCAAAATTCATAAACCCTAGACGACAA 420
DB 361 GTTTTGTATTTAACTTAACTTAACTCAAAATCAAAATTCATAAACCCTAGACGACAA 420
QY 421 CAGTCTCTTCAATATGTAAGAAAGAAAGAAAGTTTGTAGTAGCTTAAAGACACTCC 480
DB 421 CAGTCTCTTCAATATGTAAGAAAGAAAGAAAGTTTGTAGTAGCTTAAAGACACTCC 480
QY 481 ATGGAAGCTGAGTCACTGCGGCACTGCTCTCCCAATCCAAATTATAGCGATTACCA 540
DB 481 ATGGAAGCTGAGTCACTGCGGCACTGCTCTCCCAATCCAAATTATAGCGATTACCA 540
QY 541 CCGGCGACGACAAAGCTCGAGTAGCAACAACCTCTACAACTATCTGCTCCGACGC 600
DB 541 CCGGCGACGACAAAGCTCGAGTAGCAACAACCTCTACAACTATCTGCTCCGACGC 600
QY 601 AAATGGGCGGACCGTCTCTCTCCCACTTCAATTCACCTCGATCTCTCTCTCTCC 660
DB 601 AAATGGGCGGACCGTCTCTCTCTCCCACTTCAATTCACCTCGATCTCTCTCTCTCC 660
QY 661 TTTCGCAACGCGACACGCGGCACTGCTGCTCTGCGACCACTTATAGATGCTGCC 720
DB 661 TTTCGCAACGCGACACGCGGCACTGCTGCTCTGCGACCACTTATAGATGCTGCC 720
QY 721 GAAGCGCACGTCGCCATCCCATTTGATTTCTACAGGATTAAGAGCTCAAAACATTTTC 780
DB 721 GAAGCGCACGTCGCCATCCCATTTGATTTCTACAGGATTAAGAGCTCAAAACATTTTC 780
QY 781 TTAACCGATGGAATGAAGAGCATTTGAGAGTGGGTTTGAACCCGCCCAATTCGGT 840
DB 781 TTAACCGATGGAATGAAGAGCATTTGAGAGTGGGTTTGAACCCGCCCAATTCGGT 840
QY 841 TTTCGCGACGACGCTTTAATACAGCCGAGAGACGATTTCTTCAAGCTGCTTGGAAACCTTG 900
DB 841 TTTCGCGACGACGCTTTAATACAGCCGAGAGACGATTTCTTCAAGCTGCTTGGAAACCTTG 900
```

```
QY 901 TCTATCTCGGCTTAGAAGAGATACAAATGAAGTCTTGTGATGATGAAGAGCTTACA 960
DB 901 TCTATCTCGGCTTAGAAGAGATACAAATGAAGTCTTGTGATGATGAAGAGCTTACA 960
QY 961 GTCATCTGATGTTCTTGGGATTAAGTAAATTTCCATTTCCGAAATATATAAGTTTCTTC 1020
DB 961 GTCATCTGATGTTCTTGGGATTAAGTAAATTTCCATTTCCGAAATATATAAGTTTCTTC 1020
QY 1021 GTTTTAAATTTCAATGAATTTGATTAAGAGAGAACTTTATCTAGTGAAGGTTCCCTGGGGC 1080
DB 1021 GTTTTAAATTTCAATGAATTTGATTAAGAGAGAACTTTATCTAGTGAAGGTTCCCTGGGGC 1080
QY 1081 TCTCTGTGTATTTGCAAGAGGTGTGAGACTGAGATAGTTCTTCGAGTTGTGAGGCTCT 1140
DB 1081 TCTCTGTGTATTTGCAAGAGGTGTGAGACTGAGATAGTTCTTCGAGTTGTGAGGCTCT 1140
QY 1141 GCTTAAGAGAGGTGTGCTTAAGCTTTAAGCAAGATGCTTTTAACTATAGCGCTTTCG 1200
DB 1141 GCTTAAGAGAGGTGTGCTTAAGCTTTAAGCAAGATGCTTTTAACTATAGCGCTTTCG 1200
QY 1201 GTTTCGATGTCGAGGAGATGCTATGCAATTCACCTGATTTTATTACTGCTTAA 1260
DB 1201 GTTTCGATGTCGAGGAGATGCTATGCAATTCACCTGATTTTATTACTGCTTAA 1260
QY 1261 TGAATTTGTTGAGGAGCTTTGAAAGCTTTTACAGATAGTTTGAAGCTTGTGTAATTTG 1320
DB 1261 TGAATTTGTTGAGGAGCTTTGAAAGCTTTTACAGATAGTTTGAAGCTTGTGTAATTTG 1320
QY 1321 ACGAGCGTTGGCTTTAATAAGACTTCTTGATTTGATTTGATTTGATTTGATTTGATTTG 1380
DB 1321 ACGAGCGTTGGCTTTAATAAGACTTCTTGATTTGATTTGATTTGATTTGATTTGATTTG 1380
QY 1381 GAGAGAGAGAGCAAGTATGCTGACCGAGATTAACGTCACAAATTAAGATAGACTTTGA 1440
DB 1381 GAGAGAGAGAGCAAGTATGCTGACCGAGATTAACGTCACAAATTAAGATAGACTTTGA 1440
QY 1441 AGAGATCACTCCGCGTTATGCTTTGAGACTTCTGAGCTTACCGCTTGTGATGATTAAGC 1500
DB 1441 AGAGATCACTCCGCGTTATGCTTTGAGACTTCTGAGCTTACCGCTTGTGATGATTAAGC 1500
QY 1501 TCGGAAAAGACTAATAGGTTTAAAGCGGTGTCGGAATATTTTGTGCTTGTGAGAGAG 1560
DB 1501 TCGGAAAAGACTAATAGGTTTAAAGCGGTGTCGGAATATTTTGTGCTTGTGAGAGAG 1560
QY 1561 TGGAGCATCAGCTCTTGTGAGGAGTTTGAACCGTGAGAACTTTATGAATAGAGCGTTTTT 1620
DB 1561 TGGAGCATCAGCTCTTGTGAGGAGTTTGAACCGTGAGAACTTTATGAATAGAGCGTTTTT 1620
QY 1621 ACGAATGACAGCTGCTGAGAGGATTAACAGTTTAACTTTTAAATTTCTTTAGCA 1680
DB 1621 ACGAATGACAGCTGCTGAGAGGATTAACAGTTTAACTTTTAAATTTCTTTAGCA 1680
QY 1681 TGAATTAATCTTAAAGTCTTCAATTTTAAATGATTTGATGATGATGATGATGATGATGAT 1740
DB 1681 TGAATTAATCTTAAAGTCTTCAATTTTAAATGATTTGATGATGATGATGATGATGATGAT 1740
QY 1741 GCTACCCCAAGCAATATTCAGAGAGAGTCAATTTGAAGTTTACAGAGTTGACTTCTCT 1800
DB 1741 GCTACCCCAAGCAATATTCAGAGAGAGTCAATTTGAAGTTTACAGAGTTGACTTCTCT 1800
QY 1801 GTGGCTCAAGCTTTTATTTGTAAGAGCCACACTTTTACAGATCTGATTAAGCAATTC 1860
DB 1801 GTGGCTCAAGCTTTTATTTGTAAGAGCCACACTTTTACAGATCTGATTAAGCAATTC 1860
QY 1861 CAGCAACTTTAGAGAGCTTAAGATTAAGGCTTAAGATTTCCGAGATTTGATTAAGATTA 1920
DB 1861 CAGCAACTTTAGAGAGCTTAAGATTAAGGCTTAAGATTTCCGAGATTTGATTAAGATTA 1920
QY 1921 CCGAATAATTTGGAGATAGACTTCGGTCTTAGAAAGGGACTCTGTGCACTGCTTATAGGC 1980
DB 1921 CCGAATAATTTGGAGATAGACTTCGGTCTTAGAAAGGGACTCTGTGCACTGCTTATAGGC 1980
```


XX WPI: 2004-082486/08.
DR P-PSDB; ADJ38203.
XX
XX New isolated Ftn2, ARCS and/or Fzo-like nucleic acid sequences, useful
PT for further characterizing plastid division in plant cells, and in
PT varying agronomic and horticultural characteristics of economically
PT important plants.
PS
PS Claim 1; SEQ ID NO 10; 287bp; English.
XX
XX This invention relates to novel prokaryotic type or plastid division and
CC related genes and proteins. In particular, the invention relates to novel
CC Ftn2 (ARC6), ARCS and Fzo-like genes and polypeptides. The methods and
CC compositions of the present invention are useful for further
CC characterizing plastid division in plant cells, in order to vary
CC agronomic and horticultural characteristics of economically important
CC plants, such as crop, ornamental and woody plants. They can also be used
CC as herbicide targets. The present sequence is that of a DNA sequence
CC which is related to the invention.
XX
XX

Sequence 3667 BP; 982 A; 669 C; 843 G; 1173 T; 0 U; 0 Other;

Query Match 99.9%; Score 3663.8; DB 12; Length 3667;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 3665; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TGTTCGATTAAGAGAAATCAATTAAGCAATTTGCTTGATTTCAACAAGATTTG 60
DB 1 TGTTCGATTAAGAGAAATCAATTAAGCAATTTGCTTGATTTCAACAAGATTTG 60
QY 61 CTGGGCTATAGGATTCATTGAGCTGCTGTTGCTTTTACATTTACATGTCATAATAGTTTCG 120
DB 61 CTGGGCTATAGGATTCATTGAGCTGCTGTTGCTTTTACATTTACATGTCATAATAGTTTCG 120
QY 121 AATTTTACATTTTCAAGTTGATGTTAAGAAAAGAGAGGATTTGATGGGCTTTTGTGGG 180
DB 121 AATTTTACATTTTCAAGTTGATGTTAAGAAAAGAGAGGATTTGATGGGCTTTTGTGGG 180
QY 181 TTTAACTTTAAAGTATGTCAGAAATTAAGTATGTTTACTGTTCTATATAGTTAA 240
DB 181 TTTAACTTTAAAGTATGTCAGAAATTAAGTATGTTTACTGTTCTATATAGTTAA 240
QY 241 AATGAAGGCACTCCACCGGTTCTTAGTGGAATAGATTATTAGACATTTTACATCAT 300
DB 241 AATGAAGGCACTCCACCGGTTCTTAGTGGAATAGATTATTAGACATTTTACATCAT 300
QY 301 AAAGTCCGTGGGACGTGTAACATCATAGTTGTTTATTTTTCAGTACCTGTGAT 360
DB 301 AAAGTCCGTGGGACGTGTAACATCATAGTTGTTTATTTTTCAGTACCTGTGAT 360
QY 361 GTTTTTTGAATTAACCTTACTACTCAAAATCAAAATTCATAAACCCCTAGACGACAA 420
DB 361 GTTTTTTGAATTAACCTTACTACTCAAAATCAAAATTCATAAACCCCTAGACGACAA 420
QY 421 CAGTCTCTCAATATGTAATAACAGAAAGTTTGTAGTACCTTAATAAAGACACTCC 480
DB 421 CAGTCTCTCAATATGTAATAACAGAAAGTTTGTAGTACCTTAATAAAGACACTCC 480
QY 481 ATGGAAGCTCTAGTCAAGTGGCATTTGCTCTCCCAATTCAAATTAAGCGATTAACA 540
DB 481 ATGGAAGCTCTAGTCAAGTGGCATTTGCTCTCCCAATTCAAATTAAGCGATTAACA 540
QY 541 CCGGCGACGAAAGCTCGAGTAGCAACAACCTCACAACCTATGCTCGGCAAGC 600
DB 541 CCGGCGACGAAAGCTCGAGTAGCAACAACCTCACAACCTATGCTCGGCAAGC 600
QY 601 AATGGGCGACCGTCTCTCCGACTTCAATTTCACTCCGATTCCTCTCTCTCC 660
DB 601 AATGGGCGACCGTCTCTCTCCGACTTCAATTTCACTCCGATTCCTCTCTCTCC 660
QY 661 TTGCGCACCGGCAACCAACCGGCACTCTGCTCTCTGCAACATTTATGATGTC 720
DB 661 TTGCGCACCGGCAACCAACCGGCACTCTGCTCTCTGCAACATTTATGATGTC 720

DB 661 TTGCGCACCGGCAACCAACCGGCACTCTGCTCTCTGCAACATTTATGATGTC 720
QY 721 GAAGCCACGTCCTCCATCCCATATGATTTACAGGATTAAGAGCTCAACATTTTC 780
DB 721 GAAGCCACGTCCTCCATCCCATATGATTTACAGGATTAAGAGCTCAACATTTTC 780
QY 781 TTAACGATGGAATCAAGAGCAATTTGAAGCTTAGGGTTTGAACCGCGCAATTTGGT 840
DB 781 TTAACGATGGAATCAAGAGCAATTTGAAGCTTAGGGTTTGAACCGCGCAATTTGGT 840
QY 841 TTGAGGACGACGCTTTATATCAGCGGAGACAGATTCTTCAAGCTCTGCAAACTCTG 900
DB 841 TTGAGGACGACGCTTTATATCAGCGGAGACAGATTCTTCAAGCTCTGCAAACTCTG 900
QY 901 TCTATCTCGGCTCTAGAGAGAGTACATGAAGGCTCTCTGATGATGAAGAGCTACA 960
DB 901 TCTATCTCGGCTCTAGAGAGAGTACATGAAGGCTCTCTGATGATGAAGAGCTACA 960
QY 961 GTCATCACTGATGTTCTTGGGATTAAGTATTTGATTTGCGAATTAATAAGTTCTTC 1020
DB 961 GTCATCACTGATGTTCTTGGGATTAAGTATTTGATTTGCGAATTAATAAGTTCTTC 1020
QY 1021 GTTTTAAATTCATGAATTTGATTAAGAGAGAACTTTATCTAGTGAAGGTTCTCGGGGC 1080
DB 1021 GTTTTAAATTCATGAATTTGATTAAGAGAGAACTTTATCTAGTGAAGGTTCTCGGGGC 1080
QY 1081 TCTCTGTGTATATGCAAGAGGTGTGAGACTGATAGTATTTCTCGGGTTGATGAGCTCT 1140
DB 1081 TCTCTGTGTATATGCAAGAGGTGTGAGACTGATAGTATTTCTCGGGTTGATGAGCTCT 1140
QY 1141 GCTTAAGAGAGGTTGCTTAAGTCTGTTAAGCAAGATGCTGTTTATGATTAAGCGCTTGC 1200
DB 1141 GCTTAAGAGAGGTTGCTTAAGTCTGTTAAGCAAGATGCTGTTTATGATTAAGCGCTTGC 1200
QY 1201 GTTTCGATGTCGAGGAGATGCTATAGGATTTGATGATGATGATGATGATGATGATGAT 1260
DB 1201 GTTTCGATGTCGAGGAGATGCTATAGGATTTGATGATGATGATGATGATGATGATGAT 1260
QY 1261 TGAATTTGTGAGAGAACTTTGAAGCTTTTACAGATAGTTGACTTTCGTTGATATTTG 1320
DB 1261 TGAATTTGTGAGAGAACTTTGAAGCTTTTACAGATAGTTGACTTTCGTTGATATTTG 1320
QY 1321 ACGAGCGTGGCTTTATTAAGAACTTTCTGATTTGAATTAAGTATGATGATGATGATGAT 1380
DB 1321 ACGAGCGTGGCTTTATTAAGAACTTTCTGATTTGAATTAAGTATGATGATGATGATGAT 1380
QY 1381 GGAGAGAGAGCAAGTACCGCTGCAACCGGATTAACGTCACAAATTAAGATGATGATGAT 1440
DB 1381 GGAGAGAGAGCAAGTACCGCTGCAACCGGATTAACGTCACAAATTAAGATGATGATGAT 1440
QY 1441 AGAGATCACTCCGCTTATGCTTGAAGCTACTTGGCTTACCGCTTGGTATGATGATGAT 1500
DB 1441 AGAGATCACTCCGCTTATGCTTGAAGCTACTTGGCTTACCGCTTGGTATGATGATGAT 1500
QY 1501 TGGCAAAAGACTAATGCTTTAAGCGGTGTGCGGAATATTTGTGCTGTGAGAGAGG 1560
DB 1501 TGGCAAAAGACTAATGCTTTAAGCGGTGTGCGGAATATTTGTGCTGTGAGAGAGG 1560
QY 1561 TGGAGCATCAAGCTTTGTGAGGAGGTTGAACCGGGAAGAGTTTAAGATGAGAGGCTTTT 1620
DB 1561 TGGAGCATCAAGCTTTGTGAGGAGGTTGAACCGGGAAGAGTTTAAGATGAGAGGCTTTT 1620
QY 1621 ACGAATGACAGCTCTAGAGAGATACAGTTTATGATACCTTTTAAATTTCTTTAGCA 1680
DB 1621 ACGAATGACAGCTCTAGAGAGATACAGTTTATGATACCTTTTAAATTTCTTTAGCA 1680
QY 1681 TGAATATACTTTAGGTTTCTCATTTTAATGATGATGATGATGATGATGATGATGATGAT 1740
DB 1681 TGAATATACTTTAGGTTTCTCATTTTAATGATGATGATGATGATGATGATGATGATGAT 1740
QY 1741 GCTACCCCAACCATTAATTCAGAGAGTCAATTTGAAGTTTACGAAGTTGACCTTGCTCT 1800
DB 1741 GCTACCCCAACCATTAATTCAGAGAGTCAATTTGAAGTTTACGAAGTTGACCTTGCTCT 1800

```
QY 1801 GTGGCTCAAGCTTTTATTGGTAAGAACCACTTTTACAGATGCTGATTAAGCAATTC 1860
DB 1801 GTGGCTCAAGCTTTTATTGGTAAGAACCACTTTTACAGATGCTGATTAAGCAATTC 1860
QY 1861 CAGCAACTTACAGAGCTTAAGTAATGCTATGAGATTCCTCCGATGTTGTAATACA 1920
DB 1861 CAGCAACTTACAGAGCTTAAGTAATGCTATGAGATTCCTCCGATGTTGTAATACA 1920
QY 1921 CGGAATTAATGGGAGATAGACTTCGGTCTAGAAAGGGGACTGTCGATGCTGTTATAGGC 1980
DB 1921 CGGAATTAATGGGAGATAGACTTCGGTCTAGAAAGGGGACTGTCGATGCTGTTATAGGC 1980
QY 1981 AAAGTGAATGAATGCCGTAATGTTGGCTTGAACAGTGAAGATTCACATATAGCAAT 2040
DB 1981 AAAGTGAATGAATGCCGTAATGTTGGCTTGAACAGTGAAGATTCACATATAGCAAT 2040
QY 2041 CAGCTATGTTGAGAGTTTGTGTTGAGAAATTCMAATCGTATGACATATGATCTCCCT 2100
DB 2041 CAGCTATGTTGAGAGTTTGTGTTGAGAAATTCMAATCGTATGACATATGATCTCCCT 2100
QY 2101 GGAATATGCAAAATGTTGAGAACTGAGTTGGAGGGGTTGCTTCTAGGTTCAAGAGC 2160
DB 2101 GGAATATGCAAAATGTTGAGAACTGAGTTGGAGGGGTTGCTTCTAGGTTCAAGAGC 2160
QY 2161 ACCAAAGATTAATAATTTAACTCGGGGACTATATGATGATCTATGGTTTGAATTAC 2220
DB 2161 ACCAAAGATTAATAATTTAACTCGGGGACTATATGATGATCTATGGTTTGAATTAC 2220
QY 2221 TTGGAAGAGTGAAGTGAAGTTCAAGGTTCTCTTTAGCTGCTGCTGCACTATGCGAAG 2280
DB 2221 TTGGAAGAGTGAAGTGAAGTTCAAGGTTCTCTTTAGCTGCTGCTGCACTATGCGAAG 2280
QY 2281 ATTGAGCGGAGCATGTGAAGCTAGTGTATGCAAGGCACTGCAAGAAATTTTCTTCC 2340
DB 2281 ATTGAGCGGAGCATGTGAAGCTAGTGTATGCAAGGCACTGCAAGAAATTTTCTTCC 2340
QY 2341 CGCTATACAGATAGAACTCGGCTGAACCAAGATGTCAGAGAGCAAGTGTATTGATGA 2400
DB 2341 CGCTATACAGATAGAACTCGGCTGAACCAAGATGTCAGAGAGCAAGTGTATTGATGA 2400
QY 2401 GATTCCTGTTGTTAACTATGAGCCGTGATGCTGAGCTGTGTTCTTTATTGCAAGACT 2460
DB 2401 GATTCCTGTTGTTAACTATGAGCCGTGATGCTGAGCTGTGTTCTTTATTGCAAGACT 2460
QY 2461 GTAAAGCCCTCTGAATCTTTGAACTAATGATTTGCAATCGAGCTGGGGTCTCAGAG 2520
DB 2461 GTAAAGCCCTCTGAATCTTTGAACTAATGATTTGCAATCGAGCTGGGGTCTCAGAG 2520
QY 2521 AGTAGCGTTGATGAATCTACTGTTGAAATGTCGGTGTGATATGTTAAAGAGCAAGT 2580
DB 2521 AGTAGCGTTGATGAATCTACTGTTGAAATGTCGGTGTGATATGTTAAAGAGCAAGT 2580
QY 2581 GTGAAGATCTAGCTGCTGCTGTGCGAATGCACTGATTTCACTGTTACGCCAAGATAT 2640
DB 2581 GTGAAGATCTAGCTGCTGCTGTGCGAATGCACTGATTTCACTGTTACGCCAAGATAT 2640
QY 2641 TTTCTTAAAGAGCTCATCTTTTCAACGCAAGATATGCTTTCTATGGAATCTGAT 2700
DB 2641 TTTCTTAAAGAGCTCATCTTTTCAACGCAAGATATGCTTTCTATGGAATCTGAT 2700
QY 2701 GTGCGTACCATAGATATGATTAATGATGCAATTTTCAATATCTGCAATTCCTCAAAATA 2760
DB 2701 GTGCGTACCATAGATATGATTAATGATGCAATTTTCAATATCTGCAATTCCTCAAAATA 2760
QY 2761 TGCTGTTTGTGAGCTAAGAACTATGTTCCACTTAATATATGTCCTCAAAAGTTGATAC 2820
DB 2761 TGCTGTTTGTGAGCTAAGAACTATGTTCCACTTAATATATGTCCTCAAAAGTTGATAC 2820
QY 2821 AAGATTTAACAAGTGTGAGTAAATTTCACTAATATATGCTGATTTTGTGATCAAA 2880
DB 2821 AAGATTTAACAAGTGTGAGTAAATTTCACTAATATATGCTGATTTTGTGATCAAA 2880
```

```
QY 2881 CTGTAGACAGAAATGTAAATTTCACTCTCAACATTTCTGTTTAAATTAAGTAAATG 2940
DB 2881 CTGTAGACAGAAATGTAAATTTCACTCTCAACATTTCTGTTTAAATTAAGTAAATG 2940
QY 2941 AGATTGCTGATGAGCTTTGTGCAACTTTTCTTCCGTATGTTTCTTTGATTT 3000
DB 2941 AGATTGCTGATGAGCTTTGTGCAACTTTTCTTCCGTATGTTTCTTTGATTT 3000
QY 3001 AAGGCTCAGTCAAGCTGACGATTCAGAAAGCACTTCCAGAAATGATGCTAGAGACTGCA 3060
DB 3001 AAGGCTCAGTCAAGCTGACGATTCAGAAAGCACTTCCAGAAATGATGCTAGAGACTGCA 3060
QY 3061 GAATATGATTCAGATGCGCAAGATTAAGTCTCTGCTTTTGGGCTGATCACCGAT 3120
DB 3061 GAATATGATTCAGATGCGCAAGATTAAGTCTCTGCTTTTGGGCTGATCACCGAT 3120
QY 3121 AGAAATGTTACAGAGGTAAGGAAATTAATCTACAAATGTTGAAATCTGT 3180
DB 3121 AGAAATGTTACAGAGGTAAGGAAATTAATCTACAAATGTTGAAATCTGT 3180
QY 3181 TGACATGATTAATAGTCTGCTGCTGCTGTTGATTCGTATATTAATAGGTTTGAATGGC 3240
DB 3181 TGACATGATTAATAGTCTGCTGCTGCTGTTGATTCGTATATTAATAGGTTTGAATGGC 3240
QY 3241 GAATGCTGAAGATTTGACTGACAGAGCACTGAAATGCCAGCTTGCTGTTATG 3300
DB 3241 GAATGCTGAAGATTTGACTGACAGAGCACTGAAATGCCAGCTTGCTGTTATG 3300
QY 3301 ATTATACCTGTTGAAATCTATCTGTGACAGTGTGACAGTCTGACAGATGGAACCGTG 3360
DB 3301 ATTATACCTGTTGAAATCTATCTGTGACAGTGTGACAGTCTGACAGATGGAACCGTG 3360
QY 3361 CTGCTGGAAGCACTCTGAGAGAGTCTGCTGCTATCTGATTTGTTGATCCAGAA 3420
DB 3361 CTGCTGGAAGCACTCTGAGAGAGTCTGCTGCTATCTGATTTGTTGATCCAGAA 3420
QY 3421 ACAATGCTACTGATGTCAAGACTTACACAAAGATAGCAAGATTTTCTGCTCAAGTCA 3480
DB 3421 ACAATGCTACTGATGTCAAGACTTACACAAAGATAGCAAGATTTTCTGCTCAAGTCA 3480
QY 3481 GGTGAAATACATGGAAGCTCTGTTCTTGATCATATATATCATATGAGCAATGCT 3540
DB 3481 GGTGAAATACATGGAAGCTCTGTTCTTGATCATATATATCATATGAGCAATGCT 3540
QY 3541 GAGCTTGCAGATTTCTCTTGTCTGTAATTTCTCTCTAAGTTAGTGTAAATGA 3600
DB 3541 GAGCTTGCAGATTTCTCTTGTCTGTAATTTCTCTCTAAGTTAGTGTAAATGA 3600
QY 3601 ACACAAAAAATTAACGTTCTTGACACACCTTTTCTTGAATCTAATCAATTAAGG 3660
DB 3601 ACACAAAAAATTAACGTTCTTGACACACCTTTTCTTGAATCTAATCAATTAAGG 3660
QY 3661 GCTACAA 3667
DB 3661 GCTACAA 3667
```

RESULT 3
ADT14901
ID ADT14901 standard; cDNA; 2679 BP.
XX
XX ADT14901;
DT 13-JAN-2005 (first entry)
XX
DE Plant cDNA, Seq ID 227.
XX
XX Plant; ss; gene; transgenic; cold tolerance; growth rate;
XX drought tolerance; disease resistance; galactomanan production;
XX plant growth regulator; heat tolerance; herbicide tolerance;
XX lignin production; extreme osmotic condition tolerance;
XX pathogens resistance; pest resistance; yield improvement; seed oil yield;
XX seed protein yield.

XX Viridiplantae.
OS
XX
XX US2004216190-A1.
XX
XX 28-OCT-2004.
XX
XX 18-DEC-2003; 2003US-00739930.
XX
XX 28-APR-2003; 2003US-00424599.
XX
XX 28-APR-2003; 2003US-00425115.
XX
XX (KOVA/) KOVALIC D K.
XX
XX Kovalic DK;
XX
XX WPI, 2004-757369/74.
XX
XX
XX New recombinant DNA constructs useful in the field of biochemistry and
PT genetics, and in particular for producing transgenic plants with improved
PT biological characteristics.
XX
XX
XX Claim 1; SEQ ID NO 227; 14pp; English.
XX
XX The invention relates a recombinant DNA construct comprising a
CC polynucleotide having any of 5544 nucleotide sequences (CDNs SEQ ID NO:
CC 1-5544) and encoding a polypeptide with any of 5544 amino acid sequences
CC (SEQ ID NO: 5545-11088). The CDNs and proteins are from corn, soybean,
CC Arabidopsis, wheat and rape but the specification does not indicate which
CC sequences is derived from which organism. Also included is a method of
CC producing a plant having an improved property, comprising transforming a
CC plant with a recombinant DNA construct comprising a promoter region
CC functional in a plant cell operably joined to a polynucleotide encoding a
CC polypeptide associated with the property, and growing the transformed
CC plant. The property is selected from improving plant cold tolerance, for
CC manipulating growth rate in plant cells by modification of the cell cycle
CC pathway, for improving plant drought tolerance, for providing increased
CC resistance to plant diseases, for galactomanan production, for production
CC of plant growth regulators, for improving plant tolerance, for
CC improving plant tolerance to herbicides, for increasing the rate of
CC homologous recombination in plants, for lignin production, for improving
CC plant tolerance to extreme osmotic conditions, for improving plant
CC tolerance to pathogens or pests, for yield improvement by modification of
CC photosynthesis, for modifying seed oil yield and/or content, for
CC modifying seed protein yield and/or content, for yield improvement by
CC modification of carbohydrate, nitrogen or phosphorus use and/or uptake
CC and for yield improvement by providing improved plant growth and
CC development under at least one stress condition. The polynucleotide may
CC also encode a plant transcription factor. The methods and compositions of
CC the present invention are useful in the field of biochemistry and
CC genetics, in particular for producing transgenic plants with improved
CC biological characteristics such as increased yield, improved nitrogen
CC flow, increasing plant tolerance to cold or heat, improving plant
CC tolerance to extreme osmotic and drought conditions, and improving plant
CC tolerance to plant pests or pathogens. They can also be used in physical
CC arrays of molecules, plant breeding markers, computer-based storage and
CC analysis systems. The present sequence is one of the 5544 plant cDNA
CC sequences of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20040216190.
XX
XX
XX Sequence 2679 BP; 712 A; 545 C; 653 G; 769 T; 0 U; 0 Other;
XX
XX
XX Query Match 53.7%; Score 1968.4; DB 13; Length 2679;
XX Best Local Similarity 80.5%; Pred. No. 0;
XX Matches 2657; Conservative 0; Mismatches 11; Indels 632; Gaps 5;

DB 62 TTCAATATGTAAACAAACAAAGTTTGTAGTAGCTTAAAGAACACCTCCATGGAAG 121
OY |||||
OY 488 CTCTAGTCAAGTGGGATTTGTTCTCTCCCATTCCTAAATATGAGCCATACACCGCGA 547
DB 122 CTCTAGTCAAGTGGGATTTGTTCTCTCCCATTCCTAAATATGAGCCATACACCGCGA 181
OY 548 CGACAAAGCTCCGACGTAGCCAAACACCTCTACAACTATCTGCTCCGCGCAAAATGAG 607
DB 182 CGACAAAGCTCCGACGTAGCCAAACACCTCTACAACTATCTGCTCCGCGCAAAATGAG 241
OY 608 CCGACCGTCTTCTCTCCGACTTCAATTTCACTCCGATTCCTCTCTCTCTCTCTCTCT 667
DB 242 CCGACCGTCTTCTCTCCGACTTCAATTTCACTCCGATTCCTCTCTCTCTCTCTCTCT 301
OY 668 CCGACCGTCTTCTCTCCGACTTCAATTTCACTCCGATTCCTCTCTCTCTCTCTCTCT 727
DB 302 CCGACCGTCTTCTCTCCGACTTCAATTTCACTCCGATTCCTCTCTCTCTCTCTCTCT 361
OY 728 ACGTCCCATCCCATTTGATTTTCTACAGGATTTAGGAGCTCAACACATTTCTTAACG 787
DB 362 ACGTCCCATCCCATTTGATTTTCTACAGGATTTAGGAGCTCAACACATTTCTTAACG 421
OY 788 ATGGAATCAGAAAGCATTTGAAAGCTAGGTTTGGAAACCGCGCAATTTGTTTCAAG 847
DB 422 ATGGAATCAGAAAGCATTTGAAAGCTAGGTTTGGAAACCGCGCAATTTGTTTCAAG 481
OY 848 ACGAGCTTTAATCAGCGGACGAGACAGATTTCTCAGGCTGCGGAAACTCTGCTATC 907
DB 482 ACGAGCTTTAATCAGCGGACGAGACAGATTTCTCAGGCTGCGGAAACTCTGCTATC 541
OY 908 CTCGGTCTAGAAAGAGTACAAATGAAGTCTTCTTGAATGAAAGACTACATCA 967
DB 542 CTCGGTCTAGAAAGAGTACAAATGAAGTCTTCTTGAATGAAAGACTACATCA 601
OY 968 CTGATTTCTTGGGATTAAGTATTTTCATTTGGAATATTAAGTTTCTGCTTTA 1027
DB 602 CTGATTTCTTGGGATTTAAGTATTTTCATTTGGAATATTAAGTTTCTGCTTTA 618
OY 1028 TTTGATGATTTGATTAAGAAAGAACTTTATCTAGTGAAGTTCTGCGGCTCTCTGT 1087
DB 619 TTTGATGATTTGATTAAGAAAGAACTTTATCTAGTGAAGTTCTGCGGCTCTCTGT 639
OY 1088 GTATTGCAAGAAAGTGTGAGACTGAGATGTTCTGCGGCTGTGAGGCTCTGCTTAAG 1147
DB 640 GTATTGCAAGAAAGTGTGAGACTGAGATGTTCTGCGGCTGTGAGGCTCTGCTTAAG 699
OY 1148 GAGAGTTTGGCTTAAGTGTGTTAAGCAAGATGTGTTTATGTTATGGGCTTCTCTC 1207
DB 700 GAGAGTTTGGCTTAAGTGTGTTAAGCAAGATGTGTTTATGTTATGGGCTTCTCTC 759
OY 1208 GATGTTCTGAGAGGATGCTATGAGCTTGGATCCACTGATTTTATTACTGTTATGAGTT 1267
DB 760 GATGTTCTGAGAGGATGCTATGAGCTTGGATCCACTGATTTTATTACTGTTATGAGTT 819
OY 1268 GTTGAAGAAAGCTTGAAGCTTTTACAGATGTTGACTGTTGCTTTGTTTGAAGAGCG 1327
DB 820 GTTGAAGAAAGCTTGAAGCTTTTACAGATGTTGACTGTTGCTTTGTTTGAAGAGCG 844
OY 1328 TTGGCTTTATTAAGAACTTTCTGATTTGATTTGATTTGATTTGAGTCTTGTGTAGAGAA 1387
DB 845 TTGGCTTTATTAAGAACTTTCTGATTTGATTTGATTTGATTTGAGTCTTGTGTAGAGAA 852
OY 1388 GGAGCAAGTAGCTTTGACCGGATTTTACGTCACAAATTTGATGAGACTTTGGAAGAGATC 1447
DB 853 GGAGCAAGTAGCTTTGACCGGATTTTACGTCACAAATTTGATGAGACTTTGGAAGAGATC 912
OY 1448 ACTCGGCTTATGCTTGAAGCTTACCTTGAAGCTTACCTGCTTGAAGCTTGAAGAA 1507
DB 913 ACTCGGCTTATGCTTGAAGCTTACCTTGAAGCTTACCTGCTTGAAGCTTGAAGAA 972
OY 1508 AGACTAAATGTTTAAAGCGGTGCGGAATTTTGTGTGCTGTTGGAAGAGAGTGGAGCA 1567

XX ADJ38210;
AC
XX
DT 06-MAY-2004 (first entry)
XX
DE Plastid division-related Arc6 orthologue gene 3.
XX
KM prokaryotic type; plastid division; Ftn2; ARC6; ARC5; Fzo; plant cell;
KM agronomic; horticultural; crop plant; ornamental plant; woody plant;
KM herbicide target; gene; ds.
OS
XX Arabidopsis thaliana.
XX
PN W02004001003-A2.
XX
PD 31-DEC-2003.
XX
PP 20-JUN-2003; 2003WO-US019536.
XX
PR 20-JUN-2002; 2002US-0390140P.
PR 09-AUG-2002; 2002US-0402242P.
PR 20-JUN-2003; 2003US-00600070.
XX
PA (UNMS) UNIV MICHIGAN STATE.
XX
PI Oeteryoung KW, Vittha S, Koksharova OA, Gao H;
XX
DR WPI: 2004-082486/08.
DR P-PSDB: ADJ38209.
XX
PT New isolated Ftn2, ARC5 and/or Fzo-like nucleic acid sequences, useful
PT for further characterizing plastid division in plant cells, and in
PT varying agronomic and horticultural characteristics of economically
PT important plants.
XX
PS Disclosure; Fig 8; 287pp; English.
XX
XX This invention relates to novel prokaryotic type or plastid division and
CC related genes and proteins. In particular, the invention relates to novel
CC Ftn2 (ARC6), ARC5 and Fzo-like genes and polypeptides. The methods and
CC compositions of the present invention are useful for further
CC characterizing plastid division in plant cells, in order to vary
CC agronomic and horticultural characteristics of economically important
CC plants, such as crop, ornamental and woody plants. They can also be used
CC as herbicide targets. The present sequence is that of a gene which is
CC related to the invention.
XX
XX Sequence 2637 BP; 706 A; 535 C; 644 G; 752 T; 0 U; 0 Other;
Query Match 52.7%; Score 1931; DB 12; Length 2637;
Best Local Similarity 80.4%; Pred. No. 0;
Matches 2616; Conservative 0; Mismatches 5; Indels 632; Gaps 5;
QY 368 GATTAACTTACTACTCAAAATCAAAATCCATAAACCCTAGACGCAAGATCTC 427
DB 1 GATTAACTTACTACTCAAAATCAAAATCCATAAACCCTAGACGCAAGATCTC 60
QY 428 TTCAATATGTAAACAGAACAAAGTTTGTAGTAGCTTAAAGACCTCCATGAG 487
DB 61 TTCAATATGTAAACAGAACAAAGTTTGTAGTAGCTTAAAGACCTCCATGAG 120
QY 488 CTCTGAGTCACGTGGCGCATTTGCTCTCCCATTCCAATTATGCGGATTACCAACGCGGA 547
DB 121 CTCTGAGTCACGTGGCGCATTTGCTCTCCCATTCCAATTATGCGGATTACCAACGCGGA 180
QY 548 CGAAGAGCTCGAGCTGAGCAACAAACCTCTACAACTATGCTCGCGCGCAAAATGGG 607
DB 181 CGAAGAGCTCGAGCTGAGCAACAAACCTCTACAACTATGCTCGCGCGCAAAATGGG 240
QY 608 CCGACCGCTCTCTCCGAGCTTCAATTTCACTCGGATTCTCTCTCTCTCTCTGCGCA 667
DB 241 CCGACCGCTCTCTCTCCGAGCTTCAATTTCACTCGGATTCTCTCTCTCTCTCTGCGCA 300

QY 668 CCGCACCACCCACCGCACTCTGCTCTCTGCCACCATCTATTGATGTCGGAACGC 727
DB 301 CCGCACCACCCACCGCACTCTGCTCTCTGCCACCATCTATTGATGTCGGAACGC 360
QY 728 ACGTCCCATCCCATGATTTTCTACAGAGTATTAGAGGCTCAACACATTTCTTAACG 787
DB 361 ACGTCCCATCCCATGATTTTCTACAGAGTATTAGAGGCTCAACACATTTCTTAACG 420
QY 788 ATGGAATCAGAGAGCATTCGAAAGCTAGGGTTTCGAACCCGCGCAATTCGTTTCAAGC 847
DB 421 ATGGAATCAGAGAGCATTCGAAAGCTAGGGTTTCGAACCCGCGCAATTCGTTTCAAGC 480
QY 848 ACGACGCTTTAATACGCCGAGACAGATTTCTCAAGCTGCTTGGAACCTGTCTAATC 907
DB 481 ACGACGCTTTAATACGCCGAGACAGATTTCTCAAGCTGCTTGGAACCTGTCTAATC 540
QY 908 CTGGGCTAGAGAGATAGATGAAGGCTCTCTGATGATGGAAGAGCTACAGTCATCA 967
DB 541 CTGGGCTAGAGAGATAGATGAAGGCTCTCTGATGATGGAAGAGAGCTACAGTCATCA 600
QY 968 CTGATGTTCTCTGGGATTAAGTATTTGATTTGGAATTAATTAAGTTCTTGCTTTTAA 1027
DB 601 CTGATGTTCTCTGGGATTTT----- 617
QY 1028 TTTCATGAATTCGATTAAGAGAGAACTTTTATCTAGTGAAGTTCTTGCGGCTCTGT 1087
DB 618 -----AAGGTTCTCTGGGCTCTGT 638
QY 1088 GTATTGCAAGAGGTGGTGAAGTGAAGTAGTTCTTGCGGTTGGTGAAGCTTGCTTAAG 1147
DB 639 GTATTGCAAGAGGTGGTGAAGTGAAGTAGTTCTTGCGGTTGGTGAAGCTTGCTTAAG 698
QY 1148 GAGAGGTTGCTTAAGTGTGTTAAGCAAGATGATGTTTACTATAGGCTTGCTTGCTC 1207
DB 699 GAGAGGTTGCTTAAGTGTGTTAAGCAAGATGATGTTTACTATAGGCTTGCTTGCTC 758
QY 1208 GATGCTCGAGGGATGCTATGAGCATTCGATTCACCTGATTTTATCTGCTATGAGTT 1267
DB 759 GATGCTCGAGGGATGCTATGAGCATTCGATTCACCTGATTTTATCTGCTATGAGTT 818
QY 1268 GTTGAGGACCTTTGAAGCTTTTACAGGTAGTTGACTTGCTTGTTGAATTTGACGAGC 1327
DB 819 GTTGAGGACCTTTGAAGCTTTTAC----- 843
QY 1328 TTGGCTTATTAAGAACCTTCTGATTTGATATCTTGTATTGAGTCTTGTAAGAGAA 1387
DB 844 -----AGAGGAA 851
QY 1388 GAGCAAGTAGCTCTTGACCGGATTTACGTGCACAAATGATGAGACTTTGGAAGATC 1447
DB 852 GAGCAAGTAGCTCTTGACCGGATTTACGTGCACAAATGATGAGACTTTGGAAGATC 911
QY 1448 ACTCCGGCTATGCTTGAGAGCTTACCTTGCGCTTGCTGCTGATGATTAACGCTGCGAA 1507
DB 912 ACTCCGGCTATGCTTGAGAGCTTACCTTGCGCTTGCTGCTGATGATTAACGCTGCGAA 971
QY 1508 AGACTAATGCTTTAAGCGGTGCGGAATATTTGTGCTCTGTTGAGAGAGGTGAGCA 1567
DB 972 AGACTAATGCTTTAAGCGGTGCGGAATATTTGTGCTCTGTTGAGAGAGGTGAGCA 1031
QY 1568 TCAGCTTTGTGGGGGTTTGACCGCTGAGAGATTTATGAATGAGCGCTTTTACGAATG 1627
DB 1032 TCAGCTTTGTGGGGGTTTGACCGCTGAGAGATTTATGAATGAGCGCTTTTACGAATG 1091
QY 1628 ACAAGCTGAGAGAGATACAGTTAGATACCTTTTAAATTTCTTAGATGATATA 1687
DB 1092 ACAAGCTGAGAG----- 1104
QY 1688 ACTTAGGTTTCTCATTTAATGATGATGTTGTTAGGTTGATCTTTTGTAGCTAACCC 1747
DB 1105 -----AGTTGATCTTTTGTAGCTAACCC 1128
QY 1748 CAAGCAATATTCCAGCAGAGTCAATTGAAGTTTACGAAGTTGACCTTGCTTGAGCTC 1807

Db	1129	CAAGCAATATCCAGCAGATCATTTGAAAGTTTACGAAGTTGCATTGCTTGTGGCTC	1188
Qy	1808	AAGCTTTTATTTGGTAAGAGCCACACTTTTACAGATGCTGATAGCAATTCGACAAAC	1867
Db	1189	AAGCTTTTATTTGGTAAGAGCCACACTTTTACAGATGCTGATAGCAATTCGACAAAC	1248
Qy	1868	TTCCAGCAGGCTTAAGGTAAATGGCTATGAGGATTTCTCGCGATGTGTATATACAGGAATA	1927
Db	1249	TTCCAGCAGGCTTAAGGTAAATGGCTATGAGGATTTCTCGCGATGTGTATATACAGGAATA	1308
Qy	1928	ATTGGGAAATATGACTTCGGTCTAGAAAGGGGACCTGGCACTGCTTATAGCAAAAGTGT	1987
Db	1309	ATTGGGAAATATGACTTCGGTCTAGAAAGGGGACCTGTGCACTGCTTATAGGCAAAAGTGT	1368
Qy	1988	ATGATGCGGTATGTGTTTGGGCTTACAGCTGAGGATTCACAAATATAGAAATCCAGCTA	2047
Db	1369	ATGATGCGGTATGTGTTTGGGCTTACAGCTGAGGANTTCACAAATATAGAAATCCAGCTA	1428
Qy	2048	TTTGTGAGATTGTTTGGAGAAATTCMAATCGTATGACATGATGATCTCCCTGAGCTAT	2107
Db	1429	TTTGTGAGATTGTTTGGAGAAATTCMAATCGTATGACAAATGATGATCTCCCTGAGCTAT	1488
Qy	2108	GCAAAATGTTGGAAACCTGGTTGGCAGGGGTGTCTTCCATAGGTTACAGAACCCAAAG	2167
Db	1489	GCAAAATGTTGGAAACCTGGTTGGCAGGGGTGTCTTCCATAGGTTACAGAAACCAAAG	1548
Qy	2168	ATTAATAAATTTTAAACTCGGGGACTACTATGATGATCCTATGCTTTTGGATCTTGGAAA	2227
Db	1549	ATTAATAAATTTTAAACTCGGGGACTACTATGATGATCCTATGCTTTTGGATCTTGGAAA	1608
Qy	2228	GAGTGGAGTAGTTCAAGGGTTCCTCTTATGCTGTGTCGCACTATGAGCAAGATTGGAG	2287
Db	1609	GAGTGGAGTAGTTCAAGGGTTCCTCTTATGCTGTGTCGCACTATGAGCAAGATTGGAG	1668
Qy	2288	CCGAGCATGTGAAAGCTATGTGCTATGCAAGGCACTGCAAGAAATTTTCTTCCGCTATA	2347
Db	1669	CCGAGCATGTGAAAGCTATGTGCTATGCAAGGCACTGCAAGAAATTTTCTTCCGCTATA	1728
Qy	2348	CAGATAGAAACTCGGGCTGAAACCCAAAGGATGTGCAAGAGACAGTGTATAGTATAGTCTG	2407
Db	1729	CAGATAGAAACTCGGGCTGAAACCCAAAGGATGTGCAAGAGACAGTGTATAGTATAGTCTG	1788
Qy	2408	TTTGTAAACAATGTAGGCGCGTATGAGTGTGACCTGTGATCTTATATGCGAAGCTGTAGAC	2467
Db	1789	TTTGTAAACAATGTAGGCGCGTATGAGTGTGACCTGTGATCTTATATGCGAAGCTGTAGAC	1848
Qy	2468	CCTCTGAAAACCTTGAAAACCTATATGATATGCAATTCGAGCTGGGGTCTCAGAGATAGCG	2527
Db	1849	CCTCTGAAAACCTTGAAAACCTATATGATATGCAATTCGAGCTGGGGTCTCAGAGATAGCG	1908
Qy	2528	TTTGAATGAAAACCTAGCTTGGAAATGTCCGTGCTGATATGTTTAAAGAGGCAAGTGTAGAA	2587
Db	1909	TTTGAATGAAAACCTAGCTTGGAAATGTCCGTGCTGATATGTTTAAAGAGGCAAGTGTAGAA	1968
Qy	2588	TTCCTAGCTGCTGATGTGGCAATTTGAGACTGATTTCACTGTTCAAGCCAGAAAGTATTTCTTA	2647
Db	1969	TTCCTAGCTGCTGATGTGGCAATTTGAGACTGATTTCACTGTTCAAGCCAGAAAGTATTTCTTA	2028
Qy	2648	AAAGCAGCTCATCTTTTCAACGCAAGGATATGTTTCTTATGAGAACTGATGTGCTTA	2707
Db	2029	AAAGCAGCTCATCTTTTCAACGCAAGGATATGTTTCTTATGAGAACTGATGTGCTTA	2088
Qy	2708	CCATAGGTATGATTAATATGATGCAATTTTCACTATATCTGCATTGCTCAAAAATATGCTTGT	2767
Db	2089	CCA-----	2091
Qy	2768	TTTGTGAGCTAAGAACATATGTTCCCACTTATATACATGTCCCAAAAGTTGTATCCAAATTA	2827
Db	2092	-----	2091
Qy	2828	ACAAGTTGCTGAGTAAATTTCACTAATATATGCTGCTTGAATTTTTTGTATCAAACTGTAGA	2887

Db		2092	-----	----	2091
Oy	2888	CAGAAATGTAAATTCACTCTCAACATTTCTGTTTGAATAACGTAGCATTAAGATTGC			2947
Db		2092	-----	----	2091
Oy	2948	CTTAGTGTCGCTTGTGCCACTTTCTTCTCCTTGATTTTTTCTTTCCATTAGGGTCA			3007
Db		2092	-----	-TAGGGTCA	2099
Oy	3008	GTCAGACCTGCAGATTTCAGAAGCACTTCCAGATGAGTCTAGAACATCGACAGAAATATA			3067
Db	2100	GTCGAGCTGACGATTCAGAAAGCACTTCCAGATGAGTCTAGAACATCGACAGAAATATA			2159
Oy	3068	GTATCCAGAGTGGCGAAGATTAAGTCTCTGGCTTTTGGGCGCTGATCACCGCATAGAAATG			3127
Db	2160	GTATCCAGAGTGGCGAAGATTAAGTCTCTGGCTTTTGGGCGCTGATCACCGCATAGAAATG			2219
Oy	3128	TTACCAGAGGTGAGGGGAATTAATCTCAATTCAATCAATTGTGTGAAAACCTGTGGACAT			3187
Db	2220	TTACCAG-----			2226
Oy	3188	GATTAATAGTCTGAGTGCCTGTTTGATCTTGTTAATTATAGATTTTGGATGGCGAATGCT			3247
Db	2227	-----	-AGATTTTGGATGGCGAATGCT		2248
Oy	3248	GAAGATTTGGACTGCACAGACAGCTGAAACTGCGCAGCTTGCGTTGGTTTATGATTATAC			3307
Db	2249	GAAGATTTGGACTGCACAGACAGCTGAAACTGCGCAGCTTGCGTTGGTTTATGATTATAC			2308
Oy	3308	ACTGTTGAAAATATCTGTTTGAACAGTGTGACAGTCTCAGCAGATGAAACCGTGCTCTGGT			3367
Db	2309	ACTGTTGAAAATATCTGTTTGAACAGTGTGACAGTCTCAGCAGATGAAACCGTGCTCTGGT			2368
Oy	3368	GGAAGCAAATCTGGAGAGTCTGCTGTCTATCTGATTTTGGTCAATCCAGAAAAACAATGC			3427
Db	2369	GGAAGCAAATCTGGAGAGTCTGCTGTCTATCTGATTTTGGTCAATCCAGAAAAACAATGC			2428
Oy	3428	TACTGATGTGAGAACCTPACACAACAAGATCGAAGTTTCTGTGTCCAAGTCAGAGGTGAA			3487
Db	2429	TACTGATGTGAGAACCTPACACAACAAGATCGAAGTTTCTGTGTCCAAGTCAGAGGTGAA			2488
Oy	3488	AATCACTGAAAGGCTCTGTTCTTGTCATCATATATPACTCATATGTAGCATGTCTGAGCTTG			3547
Db	2489	AATCACTGAAAGGCTCTGTTCTTGTCATCATATATPACTCATATGTAGCATGTCTGAGCTTG			2548
Oy	3548	CGAGATCTCTGTTGTCGTGAAATTCCTCCTCAAGTGTGTTPAAATGAAACAACA			3607
Db	2549	CGAGATCTCTGTTGTTTGGTTPAAATTCCTCTCTPAGTTAGTGTTPAAATGAAACAACA			2608
Oy	3608	AAAATTAAAGTTTC 3620			
Db	2609	AAAATTAAAGTTTC 2621			
<hr/>					
RESULT 5					
ADJ38129 standard; cDNA; 2406 BP.					
ID	ADJ38129				
XX	ADJ38129;				
XX					
XX	06-MAY-2004 (first entry)				
DE	Arabidopsis thaliana Arc6-1 cDNA seqID1.				
XX					
KW	prokaryotic type; plastid division; Ftn2; ARC6; ARCS; Fzo; plant cell;				
KM	agronomic; horticultural; crop plant; ornamental plant; woody plant;				
RN	herbicide target; gene; ss.				
OS	Arabidopsis thaliana.				
XX					
NM	WO2004001003-A2.				

PD 31-DEC-2003.
XX
XX 20-JUN-2003; 2003WO-US019536.
XX
PR 20-JUN-2002; 2002US-0390140P.
PR 09-AUG-2002; 2002US-0402242P.
PR 20-JUN-2003; 2003US-00600070.
XX
PA (UNMS) UNITV MICHIGAN STATE.
XX
PI Oseeryoung KW, Vitna S, Kokeharova OA, Gao H;
XX WPI; 2004-082486/08.
DR P-PSDB; ADJ38202.
XX
XX New isolated Flt2, ARCS and/or Fzo-like nucleic acid sequences, useful
PT for further characterizing plastid division in plant cells, and in
PT varying agronomic and horticultural characteristics of economically
PT important plants.
XX
PS Claim 1; SEQ ID NO 1; 287bp; English.
XX
CC This invention relates to novel prokaryotic type or plastid division and
CC related genes and proteins, in particular, the invention relates to novel
CC Flt2 (ARCS), ARCS and Fzo-like genes and polypeptides. The methods and
CC compositions of the present invention are useful for further
CC characterizing plastid division in plant cells, in order to vary
CC agronomic and horticultural characteristics of economically important
CC plants, such as crop, ornamental and woody plants. They can also be used
CC as herbicide targets. The present sequence is a cDNA sequence which is
CC related to the invention.
SQ Sequence 2406 BP; 612 A; 492 C; 618 G; 684 T; 0 U; 0 Other;
Query March 47.0%; Score 1724; DB 12; Length 2406;
Best Local Similarity 79.2%; Pred. No. 0;
Matches 2406; Conservative 0; Mismatches 0; Indels 632; Gaps 5;
QY 481 ATGGAAGCTGAGTCAAGTGGGCAATGCTCTCCCATTCATATTCGCGATTACCA 540
DB 1 ATGGAAGCTGAGTCAAGTGGGCAATGCTCTCCCATTCATATTCGCGATTACCA 60
QY 541 CCGGCGACGACAAAGCTCGAGTACCAACACCTTCAACATATTCGCTCCGACG 600
DB 61 CCGGCGACGACAAAGCTCGAGTACCAACACCTTCAACATATTCGCTCCGACG 120
QY 601 AAATGGGCGGACCGCTCTCTCCGACTTCAATTCACCTCGAATCTCTCTCTCC 660
DB 121 AAATGGGCGGACCGCTCTCTCCGACTTCAATTCACCTCGAATCTCTCTCTCC 180
QY 661 TTGCGCACCGGACCAACCGGCACTCTGCTCTGCGACCATATTCATGATGATCC 720
DB 181 TTGCGCACCGGACCAACCGGCACTCTGCTCTGCGACCATATTCATGATGATCC 240
QY 721 GAAGCGCACGTCCTCCATCCCATTTGATTTCTACCAAGTATTAGAGCTCAAAACATTC 780
DB 241 GAAGCGCACGTCCTCCATCCCATTTGATTTCTACCAAGTATTAGAGCTCAAAACATTC 300
QY 781 TTAACCGGATGAATCAGAAGACATTTGAGAGTGGGTTTGAACCGCGCAATTCGGT 840
DB 301 TTAACCGGATGAATCAGAAGACATTTGAGAGTGGGTTTGAACCGCGCAATTCGGT 360
QY 841 TTCAGCGACGAGCTTTATATCAGCGGAGACAGATTTCTCAAGCTGCTTGGAAACTCTG 900
DB 361 TTCAGCGACGAGCTTTATATCAGCGGAGACAGATTTCTCAAGCTGCTTGGAAACTCTG 420
QY 901 TCTAATCTCGGTCTTGAAGAAGATGACATGAAGGTCTTCTTGATGATGAAGAAGCTACA 960
DB 421 TCTAATCTCGGTCTTGAAGAAGATGACATGAAGGTCTTCTTGATGATGAAGAAGCTACA 480
QY 961 GTCAATCAATGATGCTTGGGATPAAGGTAATTCGATTTGGAAATATAAGTTTCTTC 1020
DB 481 GTCAATCAATGATGCTTGGGAT----- 504

QY 1021 GTTTTAATTCATGAAATGGAATAAGGAAAGCACTTTATCTAGTGAAGTTCCTGGGCG 1080
DB 505 -----AAGGTTCTGGGCG 518
QY 1081 TCTCTGTGTAATTCGAGAAAGAGTGTGAGACTGAGATGATGTTCTTCGGGTGTGTAGGCTCT 1140
DB 519 TCTCTGTGTAATTCGAGAAAGAGTGTGAGACTGAGATGATGTTCTTCGGGTGTGTAGGCTCT 578
QY 1141 GCTTAAGGAGAGGTGGCTAAGTCGTTTAAGCAAGATGCTTTAGTATAGGCGCTTCG 1200
DB 579 GCTTAAGGAGAGGTGGCTAAGTCGTTTAAGCAAGATGCTTTAGTATAGGCGCTTCG 638
QY 1201 GTTTCGATGTCTCGAGGAGTGTATGCGATTCAGATTCACCTGATTTTAATTCAGTTA 1260
DB 639 GTTTCGATGTCTCGAGGAGTGTATGCGATTCAGATTCACCTGATTTTAATTCAGTTA 698
QY 1261 TGAATTTGTTGAGAGAGCTTTGAAGCTTTTACAGATGTTGACTTCGTTGGTAATTCG 1320
DB 699 TGAATTTGTTGAGAGAGCTTTGAAGCTTTTAC----- 730
QY 1321 ACAGAGCTTGGCTTTAAGAACTTCTGATTTGATTTGATTTGATGCTTGTGTA 1380
DB 731 -----A 731
QY 1381 GAGAGAAAGAGCAAGTAGCTTGCACCGAATTTACGTGCAAAATTGATGACTTTGA 1440
DB 732 GAGAGAAAGAGCAAGTAGCTTGCACCGAATTTACGTGCAAAATTGATGACTTTGA 791
QY 1441 AGAGATACCTCCGCTTATGCTTTGAGGCTACCTTGGCTTACCGCTTGTGATGATTAAGC 1500
DB 792 AGAGATACCTCCGCTTATGCTTTGAGGCTACCTTGGCTTACCGCTTGTGATGATTAAGC 851
QY 1501 TGGAAAAGACTAATGTTAAGCGGATGCGGAATATTTTGTGCTGTGAGAGAGG 1560
DB 852 TGGAAAAGACTAATGTTAAGCGGATGCGGAATATTTTGTGCTGTGAGAGAGG 911
QY 1561 TGGAGCATCACTTGTGTTGGGGTTTGACCCGTGAGAACTTTATGATAGAGCGTTT 1620
DB 912 TGGAGCATCACTTGTGTTGGGGTTTGACCCGTGAGAACTTTATGATAGAGCGTTT 971
QY 1621 ACGAATACAGCTGCTGAGAGGATACAGTTAGATCACTTTTAAATTTCTTTAGCA 1680
DB 972 ACGAATACAGCTGCTGAGG----- 991
QY 1681 TGAATATACTTAGGTTTCATTTTAATGATGTTGTGATGATGATCTTTTGTGA 1740
DB 992 -----AGGTTGATCTTTTGTGA 1008
QY 1741 GCTAACCCCAAGCAATATTCAGAGAGAGTCAATTTGAAGTTTACGAAGTTGACTTGTCT 1800
DB 1009 GCTAACCCCAAGCAATATTCAGAGAGAGTCAATTTGAAGTTTACGAAGTTGACTTGTCT 1068
QY 1801 GTGCTCAAGCTTTATTTGTTGAAGAGCCACCTTTTACAGAGATGCTGATTAAGCAATTC 1860
DB 1069 GTGCTCAAGCTTTATTTGTTGAAGAGCCACCTTTTACAGAGATGCTGATTAAGCAATTC 1128
QY 1861 CAGCAACTTACAGAGGTTAAGTATGCTATGAGAGATTCCTGAGATGTTGATGATACA 1920
DB 1129 CAGCAACTTACAGAGGTTAAGTATGCTATGAGAGATTCCTGAGATGTTGATGATACA 1188
QY 1921 CCGAATATTTGGAGATGACCTTCGCTTGAAGAGGGGACTCTGTGCACTGCTTATAGGC 1980
DB 1189 CCGAATATTTGGAGATGACCTTCGCTTGAAGAGGGGACTCTGTGCACTGCTTATAGGC 1248
QY 1981 AAAGTTATGATGATCCGTATGCTTGTGCTTGAAGATGATGATGATGATGATGATGAT 2040
DB 1249 AAAGTTATGATGATCCGTATGCTTGTGCTTGAAGATGATGATGATGATGATGATGAT 1308
QY 2041 CCAAGCTATGTTGAGAGTTGTTTGGAGAAATTCAAATGCTGATGACATGATGATGCTCT 2100
DB 1309 CCAAGCTATGTTGAGAGTTGTTTGGAGAAATTCAAATGCTGATGACATGATGATGCTCT 1368

[illegible]

Dd		2114	-----AGTTTGGATG	CC
Oy		3241	GAATGCTGAAGATTYTGACCTGCAGCAGCTGAACTCGCGACCTTGGGTGGTTATG	CC
Dd		2129	GAATGCTGAAGATTYTGACCTGCAGCAGCAGCTGAACTCGCGACCTTGGGTGGTTATG	CC
Oy		3301	ATTATTAACACTGTTGAAACTCATCTCTGTTTCACAGTGTGACAGTCTCAGCAGATGGAA	CC
Dd		2189	ATTATTAACACTGTTGAAACTCATCTCTGTTTCACAGTGTGACAGTCTCAGCAGATGGAA	CC
Oy		3361	CTCTGCTGGAAGAACAACCTGGAAGAGCTGCTGTTCTCATCTGATTTGGTTCATCCAGAA	CC
Dd		2249	CTCTGCTGGAAGAACAACCTGGAAGAGCTGCTGTTCTCATCTGATTTGGTTCATCCAGAA	CC
Oy		3421	ACAAATGCTACTGATGTGCAGAACCTCACCAACAAGATTCGAATTTCTGCTCCAGTCA	CC
Dd		2309	ACAAATGCTACTGATGTGCAGAACCTCACCAACAAGATTCGAATTTCTGCTCCAGTCA	CC
Oy		3481	GGTGGAAAATCATCTGAAGGCTCTGTTCTTGATCATATA	CC
Dd		2369	GGTGGAAAATCATCTGAAGGCTCTGTTCTTGATCATATA	CC
 RESULT 6				
ADJ38135	ID	ADJ38135	standard; cDNA; 2406 BP.	
XX	AC	ADJ38135;		
DT		06-MAY-2004	(first entry)	
XX	DE	Arabidopsis thaliana AtFtn2 CDS		
XX	KM	prokaryotic type; plastid division; Ftn2; ARC6; ARCS; Fzo; plant cell;		
XX	KM	agronomic; horticultural; crop plant; ornamental plant; woody plant;		
XX	KM	herbicide target; gene; ss.		
XX	OS	Arabidopsis thaliana.		
PN		MO2004001003-A2.		
XX	PF	31-DEC-2003.		
PD		20-JUN-2003; 2003MO-US019536..		
XX	PR	20-JUN-2002; 2002US-0390140P.		
FR		09-AUG-2002; 2002US-0402242P.		
XX	PR	20-JUN-2003; 2003US-0060070.		
PA		(UNMS) UNIV MICHIGAN STATE.		
PI		OserYoung KW, Vicha S, Koksharova OA, Gao H;		
DR		WPI; 2004-082486/08.		
DR		P-PSDB; ADJ38203.		
PT		New isolated Ftn2, ARC5 and/or Fzo-like nucleic acid sequences, useful		
PT		for further characterizing plastid division in plant cells, and in		
PT		varying agronomic and horticultural characteristics of economically		
XX		important plants.		
PS		Claim 1; SEQ ID NO 9; 287bp; English.		
CC		This invention relates to novel prokaryotic type or plastid division and		
CC		related genes and proteins. In particular, the invention relates to novel		
CC		Ftn2 (ARC6), ARC5 and Fzo-like genes and polypeptides. The methods and		
CC		compositions of the present invention are useful for further		
CC		characterising plastid division in plant cells, in order to vary		
CC		agronomic and horticultural characteristics of economically important		
CC		plants, such as crop, ornamental and woody plants. They can also be used		
CC		as herbicide targets. The present sequence is a cDNA sequence which is		
CC		related to the invention.		

XX Sequence 2406 BP; 611 A; 491 C; 619 G; 685 T; 0 U; 0 Other;
SQ Query Match 46.9%; Score 1720.8; DB 12; Length 2406;
Best Local Similarity 79.1%; Pred. No. 0; Mismatches 2404; Conservative 0; Indels 632; Gaps 5;
Matches 2404; Conservative 0; Mismatches 2; Indels 632; Gaps 5;
QY 481 ATGGAAGCTGAGTCAAGTGGGCAATGGTCTCTCCCATTCATATAGCCGATTAACA 540
DB 1 ATGGAAGCTGAGTCAAGTGGGCAATGGTCTCTCCCATTCATATAGCCGATTAACA 60
QY 541 CCGCGAGACAAAGCTCCGACGTAGCCAAACACCTTACATATATGCTCCGACAGC 600
DB 61 CCGCGAGACAAAGCTCCGACGTAGCCAAACACCTTACATATATGCTCCGACAGC 120
QY 601 AAATGGGCGGACCGCTCTCTCCGACCTTCAATTTACCTCGATTCCTCCCTCCGCTCC 660
DB 121 AAATGGGCGGACCGCTCTCTCCGACCTTCAATTTACCTCGATTCCTCCCTCCCTCC 180
QY 661 TTCCGACACCGCACCAACCAACCGCACCTCTCGTCTCTGCGACCATCTATATGATGCTCC 720
DB 181 TTCCGACACCGCACCAACCAACCGCACCTCTCGTCTCTGCGACCATCTATATGATGCTCC 240
QY 721 GAAGCGCACGTCCTCCCATCCCATTTGATTTTCAACGATTTAGAGACTCAAAACATTTTC 780
DB 241 GAAGCGCACGTCCTCCCATCCCATTTGATTTTCAACGATTTAGAGACTCAAAACATTTTC 300
QY 781 TTTAACCGATGGAATCAGAAAGACATTTGAGCTAGGGTTGAAAACCGCGCACATTTGGGT 840
DB 301 TTTAACCGATGGAATCAGAAAGACATTTGAGCTAGGGTTGAAAACCGCGCACATTTGGGT 360
QY 841 TTCAACGACGACGCTTTTAAATGACGCGAGACAGATTCTTCAAGCTGCTTGGAAAATCTTG 900
DB 361 TTCAACGACGACGCTTTTAAATGACGCGAGACAGATTCTTCAAGCTGCTTGGAAAATCTTG 420
QY 901 TCTTAATCTCGGCTCTAGAAAGAGATCAATGAGGCTCTTGTATGATGAAGAACTACA 960
DB 421 TCTTAATCTCGGCTCTAGAAAGAGATCAATGAGGCTCTTGTATGATGAAGAACTACA 480
QY 961 GTCAATCACTGANTGTTCTTGGGATPAAGTAATTCGATTTGGAAATTAAGTTTCTTC 1020
DB 481 GTCAATCACTGANTGTTCTTGGGATPAAGTAATTCGATTTGGAAATTAAGTTTCTTC 504
QY 1021 GTTTTAAATTCATGATGATTAAGAAAGAACTTTTATCTAGTGAAGTTCTCGGGC 1080
DB 505 GTTTTAAATTCATGATGATTAAGAAAGAACTTTTATCTAGTGAAGTTCTCGGGC 518
QY 1081 TCTCTGTGATTTGCAAGAGGTGCTGAGACTGAGATAGTTCTTGCGGTTGGTGAAGCTCT 1140
DB 519 TCTCTGTGATTTGCAAGAGGTGCTGAGACTGAGATAGTTCTTGCGGTTGGTGAAGCTCT 578
QY 1141 GCTTAAGAGAGGTGCTGAGAGGTGCTTAAAGCAAGATGAGTTTATGATAGGGCTTGC 1200
DB 579 GCTTAAGAGAGGTGCTGAGAGGTGCTTAAAGCAAGATGAGTTTATGATAGGGCTTGC 638
QY 1201 GTTTCCTGATGCTGAGAGGATGCTATGCAATGGATCCACTGATTTTATTAATCACTGTTA 1260
DB 639 GTTTCCTGATGCTGAGAGGATGCTATGCAATGGATCCACTGATTTTATTAATCACTGTTA 698
QY 1261 TGAAGTTTGTGAAGAGCTTTGAAGCTTTTACAGGTAAGTTGACTTGTGTAATTTG 1320
DB 699 TGAAGTTTGTGAAGAGCTTTGAAGCTTTTACAGGTAAGTTGACTTGTGTAATTTG 730
QY 1321 ACGAGCGTTGCTTTATTAAGAACTTTCTGATTTGATACCTTTGTTATGAGCTTTGTGTA 1380
DB 731 ACGAGCGTTGCTTTATTAAGAACTTTCTGATTTGATACCTTTGTTATGAGCTTTGTGTA 731
QY 1381 GGAGGAAGAGCAAGTACCTTGTGACCGGATTTTACGTGCACAAATGTATGAGACTTTGGA 1440
DB 732 GGAGGAAGAGCAAGTACCTTGTGACCGGATTTTACGTGCACAAATGTATGAGACTTTGGA 791
QY 1441 AGAGATCACTCCGCTTATGCTTGTGAGACTTACCTGCTTACCGCTTGTGATGATTAAGC 1500

DB 792 AGAGATCACTCCGCTTATGCTTGTGAGCTACTTGGCTTACCGCTTGTGATGATTAAGC 851
QY 1501 TGCGAAAAGACTAAATGTTTAAAGCGGTGTCGGAATATTTTGTGCTGTGTGGAGAGG 1560
DB 852 TGCGAAAAGACTAAATGTTTAAAGCGGTGTCGGAATATTTTGTGCTGTGTGGAGAGG 911
QY 1561 TGGAGCATCAGCTCTTGTGGGGGTTTGACCCGTGAGAGATTATGAAATGAGCGGTTTTT 1620
DB 912 TGGAGCATCAGCTCTTGTGGGGGTTTGACCCGTGAGAGATTATGAAATGAGCGGTTTTT 971
QY 1621 ACGAATGACAGCTGTCAGAGATACAGTTTAAGTAACTTTTTTAATTTCTTTAACA 1680
DB 972 ACGAATGACAGCTGTCAGAGATACAGTTTAAGTAACTTTTTTAATTTCTTTAACA 991
QY 1681 TGATATTAATTTAGTTTCTCATTTTAAATGATGATGTTGTGATGATTTGATTTTGTGA 1740
DB 992 TGATATTAATTTAGTTTCTCATTTTAAATGATGATGTTGTGATGATTTGATTTTGTGA 1008
QY 1741 GCTAACCACCAAGCAATATTTCCAGACAGTCAATTTGAAGTTTACGAAGTTGCACTTGTCTT 1800
DB 1009 GCTAACCACCAAGCAATATTTCCAGACAGTCAATTTGAAGTTTACGAAGTTGCACTTGTCTT 1068
QY 1801 GTGCTCAAGCTTTTATTTGTTAAGAACCACTTTTACAGAGATGCTGATTAAGCAATTC 1860
DB 1069 GTGCTCAAGCTTTTATTTGTTAAGAACCACTTTTACAGAGATGCTGATTAAGCAATTC 1128
QY 1861 CAGCAACTTACAGCAGCTAAGTATGCTATGAGATTCCTGCAATGTTGATATGATACA 1920
DB 1129 CAGCAACTTACAGCAGCTAAGTATGCTATGAGATTCCTGCAATGTTGATATGATACA 1188
QY 1921 CGGAATTAATGGAGATAGACTTGGCTCTAAGAAAGGGAATCTGTGCACTGCTTATAGGC 1980
DB 1189 CGGAATTAATGGAGATAGACTTGGCTCTAAGAAAGGGAATCTGTGCACTGCTTATAGGC 1248
QY 1981 AAGTTGATGAATGCGCTATGCTTGGGCTTGTGACAGATGAGATTTCAATTAAGGAAT 2040
DB 1249 AAGTTGATGAATGCGCTATGCTTGGGCTTGTGACAGATGAGATTTCAATTAAGGAAT 1308
QY 2041 CCAAGCTATGAGGATTTGTTTGGAGAAATTCGAATGATGATGATGATGATGATGATGATG 2100
DB 1309 CCAAGCTATGAGGATTTGTTTGGAGAAATTCGAATGATGATGATGATGATGATGATGATG 1368
QY 2101 GGAATATGCAAAATTTGTGAAAACCTGCTGAGAGGCTTCTTCTAGGTTTCAAGAC 2160
DB 1369 GGAATATGCAAAATTTGTGAAAACCTGCTGAGAGGCTTCTTCTAGGTTTCAAGAC 1428
QY 2161 ACCAAGATTAATAATTTTAACTCGGGCACTACATGATGATGATGATGATGATGATGATG 2220
DB 1429 ACCAAGATTAATAATTTTAACTCGGGCACTACATGATGATGATGATGATGATGATGATG 1488
QY 2221 TTGAAAAGAGTGAAGATGAGTTTCAAGGTTCTCTTTAGCTCTGCTGCAACTATGGAAGG 2280
DB 1489 TTGAAAAGAGTGAAGATGAGTTTCAAGGTTCTCTTTAGCTCTGCTGCAACTATGGAAGG 1548
QY 2281 ATTGAGCCGAGCATGTTGAAGCTATGCTATGAGGCACTGCAAGAAATTTTTCTTCC 2340
DB 1549 ATTGAGCCGAGCATGTTGAAGCTATGCTATGAGGCACTGCAAGAAATTTTTCTTCC 1608
QY 2341 CGCTATCAATTAAGAACTGCGCTGAACCCAGAGATGTCGAAGAGACAGTCTTATGAGTA 2400
DB 1609 CGCTATCAATTAAGAACTGCGCTGAACCCAGAGATGTCGAAGAGACAGTCTTATGAGTA 1668
QY 2401 GATCTGTTGGTAAACAATGATAGCCGATGATGAGCTGATGCTGATTTATGCAAGACT 2460
DB 1669 GATCTGTTGGTAAACAATGATAGCCGATGATGAGCTGATGCTGATTTATGCAAGACT 1728
QY 2461 GTAAGACCTCTGAAAACCTTTGAAAATTAATGATTAATGCAATTTGAGCTGGGCTCTAGAG 2520
DB 1729 GTAAGACCTCTGAAAACCTTTGAAAATTAATGATTAATGCAATTTGAGCTGGGCTCTAGAG 1788
QY 2521 AGTAGCGTTGATGAATACTACGTTGGAATGTCGTTGCTGATATGTTAAAGAGGCAAGT 2580
DB 1789 AGTAGCGTTGATGAATACTACGTTGGAATGTCGTTGCTGATATGTTAAAGAGGCAAGT 1848

0y 841 TTACGCGACGAGCTTTAATACGCCGAGACAGATTTCTTCAAGCTGCTTGGCAAACCTGCG 900
Db 361 TTACGCGACGAGCTTTAATACGCCGAGACAGATTTCTTCAAGCTGCTTGGCAAACCTGCG 420
0y 901 TCTAATCTCGGCTCTAGAGAGAGATCAATGAAAGTCTTCTTGATGATGAAGAACTAC 960
Db 421 TCTAATCTCGGCTCTAGAGAGAGATCAATGAAAGTCTTCTTGATGATGAAGAACTAC 480
0y 961 GTATCATCTAGTCTTGGGATAAGTAATTTGATTTGCGAATTAATAAGTTCTTTC 1020
Db 481 GTATCATCTAGTCTTGGGATAAGT----- 504
0y 1021 GTTTAATTTGATGATGATTAAGAAAGAACTTTATCTAGTGAAGTTCTCGGCGC 1080
Db 505 -----AAGTTCTCGGCTGC 518
0y 1081 TCTCTGTATTTGCAAGAAAGTGTGAGACTGAGATGTTCTTGGGTTGGAGGCTCT 1140
Db 519 TCTCTGTATTTGCAAGAAAGTGTGAGACTGAGATGTTCTTGGGTTGGAGGCTCT 578
0y 1141 GCTTAAGAGAGGTTGCCCTAAGTCGTTTAAGCAAGATGTTTATGATGCGCTTGC 1200
Db 579 GCTTAAGAGAGGTTGCCCTAAGTCGTTTAAGCAAGATGTTTATGATGCGCTTGC 638
0y 1201 GTTTCGATGTCGAGGAGATGCTATGSCATTTGATCACTGATTTTATTAATCTGTTA 1260
Db 639 GTTTCGATGTCGAGGAGATGCTATGSCATTTGATCACTGATTTTATTAATCTGTTA 698
0y 1261 TGAAGTTTGAAGAGCTTTGAAGCTTTACAGATGTTGACTTGGCTTGGTTATTTG 1320
Db 699 TGAAGTTTGAAGAGCTTTGAAGCTTTAC----- 730
0y 1321 ACGAGCTGTGGCTTTAAGAACTTTCTTGATTTGATACTTTGTATTAGACTTGTGTA 1380
Db 731 -----A 731
0y 1381 GGAGAGAGAGCAAGTACGCTTGCACCGGATTTACGTGCACAAATGATGAGACTTTGGA 1440
Db 732 GGAGAGAGAGCAAGTACGCTTGCACCGGATTTACGTGCACAAATGATGAGACTTTGGA 791
0y 1441 AGAGTCACTCCGCTTATGCTTGGAGCTTACCTTACCGCTTGGTGTATGATTAAGC 1500
Db 792 AGAGTCACTCCGCTTATGCTTGGAGCTTACCTTACCGCTTGGTGTATGATTAAGC 851
0y 1501 TGCGAAAAGACTAAATGCTTAAAGCGTGTGCGAAATATTTTGTGTCTGTGAGAGAG 1560
Db 852 TGCGAAAAGACTAAATGCTTAAAGCGTGTGCGAAATATTTTGTGTCTGTGAGAGAG 911
0y 1561 TGAGCATCAGCTCTTGTGGGGGTTTGAACCGGTGAAGAGTTTATGAATGAGCGTTTTT 1620
Db 912 TGAGCATCAGCTCTTGTGGGGGTTTGAACCGGTGAAGAGTTTATGAATGAGCGTTTTT 971
0y 1621 ACGAATGACAGTGTGAGCAAGTATACAGTTTATACCTTTTATTAATTTCTTTAGCA 1680
Db 972 ACGAATGACAGTGTGAGCA----- 991
0y 1681 TGAATTAATCTTAGGTTTCTCATTTTAATGATGTTGTGTGAGTGTGATCTTTTGTGA 1740
Db 992 -----AGTTGATCTTTTGTGA 1008
0y 1741 GCTACCCCAAGCAATATTCACAGAGATCATTTGAAGTTTGAAGAGTTGCACTTGCTCTT 1800
Db 1009 GCTACCCCAAGCAATATTCACAGAGATCATTTGAAGTTTGAAGAGTTGCACTTGCTCTT 1068
0y 1801 GTGGCTCAAGCTTTATTTGTTAAGAGCAACCTTTTACAGAGATGCTGATTAAGCAATTC 1860
Db 1069 GTGGCTCAAGCTTTATTTGTTAAGAGCAACCTTTTACAGAGATGCTGATTAAGCAATTC 1128
0y 1861 CAGCAACTTCAGCAGGCTTAAGGTAAAGCTATGAGATTTCCGCGATGTTGATATACA 1920
Db 1129 CAGCAACTTCAGCAGGCTTAAGGTAAAGCTATGAGATTTCCGCGATGTTGATATACA 1188
0y 1921 CGGAATTAATTGGAGATGATCTTCGCTCTAGAAAGGGAACCTCTGTGCACTGCTTATAGGC 1980

Db 1189 CGGAATTAATTGGAGATGATCTTCGCTCTAGAAAGGGAACCTGTGCACTGCTATATAGGC 1248
0y 1981 AAAGTTGATGATAGCCGATATGCTTGGGCTTTGACAGTGAAGGATTTGCAATATATAGAAAT 2040
Db 1249 AAAGTTGATGATAGCCGATATGCTTGGGCTTTGACAGTGAAGGATTTGCAATATATAGAAAT 1308
0y 2041 CCAGCTATTTGAGATTTGTTTGGAGAAATTCAAATGATGATGACAAATGATCTTCTCT 2100
Db 1309 CCAGCTATTTGAGATTTGTTTGGAGAAATTCAAATGATGATGATGATCTTCTCTCT 1368
0y 2101 GGACTATGCAAAATTTGTTGGAACCTGCTGGCAGAGGTTGTCTTCTTCTAGTTCAAGAC 2160
Db 1369 GGACTATGCAAAATTTGTTGGAACCTGCTGGCAGAGGTTGTCTTCTTCTAGTTCAAGAC 1428
0y 2161 ACCAAAGATTAATAATTTTAACTCGGGAACCTATGATGATCCCTATGTTTGTAGTTAC 2220
Db 1429 ACCAAAGATTAATAATTTTAACTCGGGAACCTATGATGATCCCTATGTTTGTAGTTAC 1488
0y 2221 TTGAAAAGTGAAGTATTCAGGGTTCTCTTTAGCTGCTGCTGCAACTATGCGAAG 2280
Db 1489 TTGAAAAGTGAAGTATTCAGGGTTCTCTTTAGCTGCTGCTGCAACTATGCGAAG 1548
0y 2281 ATTGAGCCGAGACTGTGAAGCTATGCTATGAGGCACTGCGAAGAGTTTCTCTTCC 2340
Db 1549 ATTGAGCCGAGACTGTGAAGCTATGCTATGAGGCACTGCGAAGAGTTTCTCTTCC 1608
0y 2341 CGCTATACAGTAACTCGGCTGAACCCAGGATGCGAAGACAGTGTATAGTTA 2400
Db 1609 CGCTATACAGTAACTCGGCTGAACCCAGGATGCGAAGACAGTGTATAGTTA 1668
0y 2401 GATCCTGTGTGAACAAATGAGGCGGTGATGAGGCTGCTGCTCTTATTTGCAAGACT 2460
Db 1669 GATCCTGTGTGAACAAATGAGGCGGTGATGAGGCTGCTGCTCTTATTTGCAAGACT 1728
0y 2461 GTTAAGACCTCTGAAAACCTTTGAACCTAATGATTAATGCAATTTGAGCTGGGCTTCAAG 2520
Db 1729 GTTAAGACCTCTGAAAACCTTTGAACCTAATGATTAATGCAATTTGAGCTGGGCTTCAAG 1788
0y 2521 AGTAGGTTGATGAACCTACGTTGAATGTCGTTGCTGATATGTTAAAGAGGCAAGT 2580
Db 1789 AGTAGGTTGATGAACCTACGTTGAATGTCGTTGCTGATATGTTAAAGAGGCAAGT 1848
0y 2581 GTGAAGATCTAGCTGTGCTGTGCAATTGCACTGATTTCACTGTTCAAGCAGAAATAT 2640
Db 1849 GTGAAGATCTAGCTGTGCTGTGCAATTGCACTGATTTCACTGTTCAAGCAGAAATAT 1908
0y 2641 TTTCTTAAAGCAGCTCATCTTTTGAACGCAAGATATGCTTCTTATGGAATCTGAT 2700
Db 1909 TTTCTTAAAGCAGCTCATCTTTTGAACGCAAGATATGCTTCTTATGGAATCTGAT 1968
0y 2701 GTGCTTACCAATAGTATGATTAATGATGCAATTTTCAATATCTGATTCCTCAATAATA 2760
Db 1969 GTGCTTACCA----- 1978
0y 2761 TGCTTGTTTGTGAGCTTAAGAACATAGTCCCACTTAATACATGTCCCAAAAGTTGATCC 2820
Db 1979 ----- 1978
0y 2821 AAGATTAACAAGTTGCTGAGTAAATTTCACTAATATAGCTGCTTGAATTTTGTGATCAA 2880
Db 1979 ----- 1978
0y 2881 CTGTAGACAGAAATGTAAATTTCACTCAACATTTCTGTTTGAATPAAGTAGATTAAG 2940
Db 1979 ----- 1978
0y 2941 AAGATTGCTTAGTGTGCTTGTGCCAATTTTCTTCTGATTTTCTTTTTCGATTT 3000
Db 1979 -----T 1979
0y 3001 AGGTCAGTCAAGACTGACATTCAGAGACACTTCCAGAAATGATGCTAGAGACTGACAGA 3060

PR 10-DEC-1999; 99US-0170255P.
XX
PA (FINC/) FINCHER K L.
XX
PI Fincher KL,
XX
XX
DR MPI; 2004-479807/45.
XX
XX New substantially purified nucleic acid molecule that encodes a cotton
PT protein or its fragment, useful as molecular tool for the targeting and
PT isolation of novel genes for plant protection and improvement.
XX
PS Claim 1, SEQ ID NO 2154; 30pp; English.

The invention relates to a substantially purified nucleic acid molecule that encodes a cotton protein or its fragment comprising an EST (expressed sequence tag) appearing as [ADP90990-ADP95919](#). Also included are a substantially purified cotton protein or its fragment encoded by a nucleic acid molecule above and a transformed plant (having a nucleic acid molecule which comprises: an exogenous promoter region which functions in a plant cell to cause the production of a mRNA molecule; a structural nucleic acid molecule comprising one of the ESTs or their complements; a 3' non-translated sequence that functions in the plant cell to cause termination of transcription and addition of polyadenylated ribonucleotides to a 3' end of the mRNA molecule). The ESTs are useful as molecular tool for the targeting and isolation of novel genes for plant protection and improvement. The ESTs are useful for developing new strategies for understanding critical plant developmental and metabolic pathways, for isolating genes and promoters, for identifying and mapping the genes involved in developmental and metabolic pathways, and for determining gene function. The cotton nucleic acid molecules are useful as molecular tags to isolate genetic regions, isolate genes, map genes, and determine gene function. The nucleic acid molecules are useful for determining if genes are members of a particular gene family and for use in marker-assisted breeding programs. The present sequence is one of the 4930 cotton ESTs of the invention. NOTE: The sequences are not displayed in the specification and are available in electronic format from the uspto.atseqdata.uspto.gov/sequence.html?PDocID=20040123338.

SQ Sequence 439 BP; 87 A; 153 C; 79 G; 119 T; 0 U; 1 Other;

Query Match	4.8%	Score 175	DB 12	Length 439
-------------	------	-----------	-------	------------

Matches 278; Conservative 0; Mismatches 141; Indels 9; Gaps 1.

QY	499	GTCCGCAATTGGTCTCTCCCAATTCCAAATTAATGCCGATTACACCGGCGACGACGAAGAAAGTTC	558
Db	12	GTTTACATCTGGGCTGTGCACTCCGACACTCTGCTCTTTATTTGCAACCTGTGTAATACCTCTCC	71
QY	559	CGACGTAAGCAACAAACCGTATACAACTATCTGCTCCGCGACGACAAATGGGCGCGACGCTTT	618
Db	72	AAATCTCCACCAACCGTTCAACTACGCTGTCTGCTCCGCGACGTAATATGGGCGCGACGCTTC	131
QY	619	CTCTCCGACTTCAATTTGACCTCG-----ATTCTCTCTCTCTCTCTCTGCGACCC	669
Db	132	CTTGGTGACTTCGAATTTCTCCGTGCCCCCGATTAATTCGCTCTCTCTCTCTCTCTCTCTCA	191
QY	670	GGCACCACCAACGCGACACTCTGCTCTCTGTGCACTATATGATGTCGCCAAGCCGAC	729
Db	192	ACCGGCACCTCTTCTCCGCCCTTACCTCTCCCTTCTCTGCCCCCTCTCTCCCGAAGCCGAC	251
QY	730	GTCCCATTCGCCAATGATTTTCTACCAAGTATTTAGAGCTCAAAACAATTTCTTAAACGAT	789
Db	252	GTTTTCATTTCCCTTGATTTTCTACAAAGTTTATAGAACCGAGACTATTTCTTAGGTGAT	311
QY	790	GGATTCAGAAAGCAATTCGAAGCTTAGGGTTTGGAAACCGCGCAATTCGGTTTCAGCGAC	849
Db	312	GGAAATTTAGAGAGCCTTAGAAGCAAGGGTTTGGAAACCGCCTCAATATGGGTTCACTCAA	371
QY	850	GACGCTTAAATAGCGCGAGACAGATTTCTCAACCTCTGCGCAAACTCTGCTATATCT	909
Db	372	GAACCATATATATAGCGGAAGACAGATTTCTTACTCTCTGCTGTGAACCCCTATATTAACCTT	431

QY	910	CGGTCTAG	917
Db	432	GGCTCTAG	439

RESULT 14
ADJ38255
ID ADJ38255 standard; cDNA; 545 BP.

AC ADJ38255;

DT 06-MAY-2004 (first entry)

DE Plastid division-related Arc6 orthologue cDNA 42.

KW prokaryotic type; plastid division; Ftn2; ARC6; ARC5; Fzo; plant cell;

KW herbicide target; gene; ss.

OS *Gossypium arboreum.*

PN WO2004001003-A2

PD 31-DEC-2003.

PF 20-JUN-2003; 2003WO-US019536.

PR 20-JUN-2002; 2002US-0390140P.

PR 20-JUN-2003; 2003US-00600070.

PA (UNMS) UNIV MICHIGAN STATE.

PI Osteryoung KW, Vitha S, Koksharova OA, Gao H;

DR WPI; 2004-082486/08.

PT New isolated Ftn2, *A*

PT varying agronomic and horticultural characteristics of economically

XIX

XXXXXXXXXXXXXXXXXXXX

SQ Sequence 545 BP; 164 A; 107 C; 120 G; 151 T; 0 U; 3 Other;

Query Match 4.4%; Score 161; DB 12; Length 545;

Matches 218; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

Oy 3224 ATATGGTTTGGATGAGCGAATGCTGAACATTTGGAAGTGCAGAGCAGCTGAACATGCGCA 328
 Db 82 AGAGGTTCTGATGCTCAATATGTTGAAGCATGAGCAGATCGTCAGCCGAAATCGCTCA 141

Oy 3384 GCTGGGGTGGTTATGATTATACACGTTGAACTATCTGTGACAGGTGACAGTCTC 3344
 Db 142 GCTTGGTTGGGTATGATATATGTCTACTGAACATGGCCATTCGACAGGTCTACCTTTC 201

3344 AGCAGATGGCAACCGTCGTCGTAAGGCAACTCTGAGAGAGCTGCTGTGCTATCTGA 3400
 202 ACTAGATGGCCACAGCAGCTGTAGTCGAAGCTACTCTGGGAAGATCCACCTGCTTGACTGA 261

```
Db      262  TGTTCATATCCGGAGAACATGCTCTAAATGTAAACTCTACACACGAGATATGAGAT 321
Oy      3464  TTTCGGTCCAGTAGGGGTGGAAATCATCGAAGGCTCTGTTCTTGATCATATATAC 3523
Db      322  GTCTGTTCCACTCAGGCTGGAAATCATCGAAGGATCTGTCTAACAATCTTAATG 381
Oy      3524  TCATATGTAGCAT 3536
Db      382  ATGTATAAGCAT 394

RESULT 15
ACN4855/C
ID      ACN4855 standard; cDNA; 552 BP.
XX
AC      ACN4855;
XX
DT      02-DEC-2004 (first entry)
XX
DE      Cotton primed seed EST Clone ID: LIB3825-027-Q6-N6-H1, SEQ:3636.
XX
KM      Cotton, plant; EST; expressed sequence tag; transgenic plant; seed;
KM      variety DP50B; library LIB3825; molecular tag; molecular marker;
KM      genetic mapping; molecular mapping; seed germination; plant growth;
KM      plant quality; plant yield; plant breeding; tissue printing; ss.
XX
OS      Gossypium hirsutum.
XX
PN      US2004123340-A1.
XX
PD      24-JUN-2004.
XX
PF      12-DEC-2001; 2001US-00021323.
XX
PR      14-DEC-2000; 2000US-0255619P.
XX
PA      (DEIK/) DEIKMAN J.
PA      (FENG/) FENG P C C.
PA      (FINCH/) FINCHER K L.
PA      (ZIEG/) ZIEGLER T E.
XX
PI      Deikman J, Feng PCC, Fincher KL, Ziegler TE;
XX
DR      WPI; 2004-479808/45.
XX
PT      New isolated nucleic acid molecule that encodes a plant protein or its
PT      fragment, useful for isolating a variety of agronomically significant
PT      genes associated with plant growth, quality or yield, and as molecular
PT      tags to map genes.
XX
PS      Claim 1; SEQ ID NO 3636; 34p; English.
XX
CC      The invention relates to 17880 cotton expressed sequence tags (ESTs;
CC      ACN45220-ACN63099). The ESTs were isolated from cDNA libraries generated
CC      from primed or non-primed seeds from variety DP50B, mature seeds from
CC      variety Coker 312 Boswell 96 field, and androecium tissue, gynoecium
CC      tissue, developing fibres, carpel walls and septa from variety
CC      Nuoton33B. The invention also relates to substantially purified
CC      proteins or their fragments encoded by nucleic acid molecules of the
CC      invention, and to transformed plants having a nucleic acid molecule
CC      comprising a nucleic acid of the invention. The cotton ESTs are useful as
CC      molecular tags to isolate genetic regions, to isolate genes, to map
CC      genes, to determine gene function and to determining whether genes are
CC      members of a particular gene family. The nucleic acid molecules may be
CC      used for isolating a variety of agronomically significant genes
CC      associated with plant growth, quality, yield, and could also serve as
CC      links in metabolic and catabolic pathways. The nucleic acid molecules are
CC      also useful for identifying genes important in initiating and maintaining
CC      seed germination or that may be used to mitigate stresses encountered
CC      during seed germination. The ESTs additionally enable the acquisition of
CC      promoters and cis-regulatory elements which will be useful to express
CC      agronomically significant genes in these tissues and/or other tissues,
```

```
CC and also permits the acquisition of molecular markers useful in breeding
CC schemes, genetic and molecular mapping, and in cloning of agronomically
CC significant genes. The nucleic acid molecules are further useful for
CC detecting the expression level or pattern of a protein or mRNA and for
CC detecting the presence or quantity of a protein by tissue printing. The
CC present sequence represents a specifically claimed BSI isolated from a
CC cotton variety DP50B primed seed cDNA library (LIB3825). The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from the US patent office at
CC seqdata.uspco.gov/sequence.html?docid=US20040123340
XX
```

Sequence 552 BP; 153 A; 122 C; 110 G; 167 T; 0 U; 0 Other;

Query Match 4.3%; Score 159.4; DB 13; length 552;

Best Local Similarity 69.3%; Pred. No. 1.2e-30;

Matches 217; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

```
Oy      3224  ATAGGTTTGGATGGCGGAATGCTGAAGATTGGACTGACAGACGAGCTGAATCGGCA 3283
Db      444  AGAGGTTCTGGATGGTCMAATGTTGAAGACATGACAGATCGACCGAATCGCTCA 385
Oy      3284  GCTTGGGTGGTTTATGATTATACACTGTGAACCTATCTGTGACAGTGAACAGTCTC 3343
Db      384  GCTTGGGTGGTATATATAATAGCTTACTGAAACATGGCCATTGACGTGTACCTTTT 325
Oy      3344  AGCAGATGAACCCGTCCTCTGTGGAAGCACTCTGGAAGAGTCTGCTTATCTGA 3403
Db      324  ACTAGATGGCGAGGAGCTGTAGTGAAGCTACTCTGGAAGAAATCCACTGCTTGACTGA 265
Oy      3404  TTTCGTTATCCAGAAAACAATGCTACTGATGTGACAACTTACACACAAGATACGAAGT 3463
Db      264  TGTTCATCATCCCGAGAACATGCTCTAATGTAAATCTCTTACACCAAGAAATATGAGAT 205
Oy      3464  TTTCGGTCCAGTCAGGGTGGAAAAATCAGTGAAGGCTGTCTTGTGATCATATATATAC 3523
Db      204  GTCTTGTTCAACTCAGGCTGGAAAAATCAGTGAAGGATCTGTCTAACAATCTTAGCTATG 145
Oy      3524  TCATATGTAGCAT 3536
Db      144  ATGTATAAGCAT 132
```

Search completed: December 10, 2005, 19:22:07
Job time : 2104.73 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 10, 2005, 16:57:24 ; Search time 18153.9 Seconds
(without alignments)
11482.131 Million cell updates/sec

Title: US-10-600-070B-3

Perfect score: 3667
1 tgcctgcattcaagagagaat.....ctatacataagggctacaa 3667

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenBdb1:*
1: gb_da:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_srs:*
11: gb_sy:*
12: gb_un:*
13: gb_vl:*
14: gb_hcg:*
15: gb_pl:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3664	99.9	3668	15	AY221468 Arabidops
2	3660.8	99.8	3664	15	AY221467 Arabidops
3	3659	99.8	85791	15	AB016888 Arabidops
4	1931	52.7	2637	15	AY091075 Arabidops
5	1752.8	47.8	2438	15	AY221469 Arabidops
6	1744.6	47.8	2436	15	AY150490 Arabidops
7	338.6	9.2	133779	14	AC158210 Medicago
8	317.6	8.7	110000	15	AP008208_011
9	317.6	8.7	150462	15	AP004885 Oryza sat
10	246.8	6.7	184752	14	AC155597 Zea mays
11	170.4	4.6	550	10	D0022571 Beta vulg
12	146.2	4.0	117818	14	AC160013 Medicago
13	57.6	1.6	7218	6	I66494
14	54	1.5	1141	6	ARS79680 Sequence
15	54	1.5	1141	6	AX083744 Sequence
16	54	1.5	250029	2	AB014820 Plasmodi
17	53.8	1.5	205796	14	AL713891 Mus muscu
18	53.8	1.5	215524	9	AL603913 Mouse DNA

19	53.8	1.5	250078	2	AE014829	AB014829 Plasmodi
20	53.8	1.5	333321	2	AC116986	AC116986 Dictyoste
21	53.2	1.5	198573	14	AC110118	AC110118 Rattus no
22	53	1.4	184039	9	AC139376	AC139376 Mus muscu
23	53	1.4	232373	14	AC134805	AC134805 Rattus no
24	52.8	1.4	617	10	BV331985	BV331985 S230P6420
25	52.6	1.4	222366	14	AC119799	AC119799 Pan trogl
26	52.4	1.4	234008	14	CR392350	CR392350 Dabo rer
27	52.4	1.4	264451	14	AC158594	AC158594 Mus muscu
28	52.2	1.4	20079	1	AE017244_09	Continuation (10 o
29	52.2	1.4	97405	1	AE017243_8	Continuation (9 of
30	52.2	1.4	110000	1	AE017244_08	Continuation (9 of
31	52	1.4	180084	9	AC140372	AC140372 Mus muscu
32	52	1.4	245924	5	BX004773	BX004773 Zebrafish
33	52	1.4	255036	9	AC118206	AC118206 Mus muscu
34	52	1.4	268597	14	AC094787	AC094787 Rattus no
35	51.6	1.4	2000	6	AX655393	AX655393 Sequence
36	51.6	1.4	23213	8	AL590436	AL590436 Human DNA
37	51.6	1.4	110000	14	AC108609_2	Continuation (3 of
38	51.6	1.4	292533	14	AC107260	AC107260 Rattus no
39	51.4	1.4	135256	9	AC163613	AC163613 Mus muscu
40	51.4	1.4	139995	9	AC131797	AC131797 Mus muscu
41	51.4	1.4	150579	9	AC126937	AC126937 Mus muscu
42	51.4	1.4	171307	8	AC104020	AC104020 Homo sapi
43	51.4	1.4	193634	14	AC068495	AC068495 Mus muscu
44	51.4	1.4	249854	14	CT009658	CT009658 Mus muscu
45	51.4	1.4	253094	14	AC113774	AC113774 Rattus no

ALIGNMENTS

RESULT 1	AY221468	3668 bp	DNA	linear	PLN 05-AUG-2003
LOCUS	AY221468				
DEFINITION	Arabidopsis thaliana division protein (ARCE) gene, complete cds.				
ACCESSION	AY221468				
VERSION	AY221468.1	GI:33436338			
KEYWORDS	Arabidopsis thaliana (thale cress)				
SOURCE	Arabidopsis thaliana				
ORGANISM	Arabidopsis thaliana (Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi				
REFERENCE	1 (bases 1 to 3668)				
AUTHORS	Vitsha,S., Froehlich,J.F., Koksharova,O., Pyke,K.A., Van Erp,H. and Oosteryoung,K.W.				
TITLE	ARCE is a J-Domain Placoid Division Protein and an Evolutionary				
JOURNAL	Descendant of the Cyanobacterial Cell Division Protein Ftn2				
PUBMED	Plant Cell 15 (6), 1918-1933 (2003)				
REFERENCES	12897262				
AUTHORS	2 (bases 1 to 3668)				
TITLE	Vitsha,S., Koksharova,O., van Erp,H., Froehlich,J.E. and Oosteryoung,K.W.				
JOURNAL	Direct Submission				
REFERENCES	Submitted (17-JAN-2003) Department of Plant Biology, Michigan State University, 166 Plant Biology Building, East Lansing, MI 48824, USA				
AUTHORS	Location/Qualifiers				
FEATURES	1..3668				
SOURCE	/organism="Arabidopsis thaliana"				
gene	/mol type="genomic DNA"				
gene	/db xref="taxon:3702"				
gene	/ecotype="Wassilewskija"				
gene	478..3515				
gene	/gene="ARCE"				
gene	/note="synonym: At5g42480"				
gene	join(<478..984,1067..1291,1379..1639,1723..2710,				
gene	3000..3133,3225..3515)				
gene	/gene="ARCE"				
gene	/product="division protein"				
gene	join(478..984,1067..1291,1379..1639,1723..2710,3000..3133,				
gene	3225..3515)				
CDS					

/gene="ARC6"
/note="inner envelope membrane-located; similar to
cyanobacterial cell division protein sl10169 of
Synchocystis; contains J-domain"
/codon_start=1
/product="division protein"
/protein_id="AA018645.1"
/db_xref="GI:33436339"
/translation="MEALSHVIGIGLSPQLRLPPTTKLRSHNTSTTICASAKMAD
RLSDFNFTSDSSSSFPATATTTATVATLPSIDRPEBHVPIDPEVOLOAORHPLT
GIRAPAPRVSKPQPGSPDALSIRPOLIQAACETLSNREYVGLLDEBPT
VITTPMDKRVGALCVLOEGSETEVLREVGALLKERLPSKPKQDVUVALALDYS
RDAAALDPDPITIGTGEVEEALKLQEBGASLAPDLQAQIDETLEETTPRYVELLG
LPAGDYAAKRLNGLSVRNLTIVSGGASALVGLTRERKNEBAFLRMAAOVDL
FVATPSNIAPASFVEYVALVAQAFIKGPHLLODADQFOGLOQAKVMAEIVL
LYDRNWEIDFGEERGLCALLIGKDECKRMGLDSDQYRNPVSLYERVVGSPDL
NDLRLGLCKLETMLAGVFRPRDPTDRTDRNSAPKDVQETVFSVDPRVNGRD
AAATMARIGAEHYKASAMOLAKVFPSTRYDNRNSAPKDVQETVFSVDPRVNGRD
GEPEVFLAEAVRSENEFTNDYAIRAGVSSSVDETTVEMSVADMLKEASVKILAGV
AIGLISLPSQRYPLKSSSSFORKDWSSMESDVATIGSRVADDBEALPRMDARTAEI
VSKWOKISLAFGPDHRIEMLPVLDGMLKIETDRAETDQLGLVDVYTLKLSVS
VTVSADGTRALVEATLEESACLSDLVHPENNATDVRYTTRVYEFWMSGKIKTEGSV
LAS"

ORIGIN

Query Match 99.9%; Score 3664; DB 15; Length 3668;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 3664; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 4 TCTGCATTAAAGCAATATCAATTATTAAGCAATTTGCTGATTTCAACAAGATTGCTT 63
Db 1 TCTGCATTAAAGCAATATCAATTATTAAGCAATTTGCTGATTTCAACAAGATTGCTT 60
OY 64 GGCATAGGATTCATTGGCTCTGTTGCTTTTACATTACATGTCATATAGTTTGCAT 123
Db 61 GGCATAGGATTCATTGGCTCTGTTGCTTTTACATTACATGTCATATAGTTTGCAT 120
OY 124 TTTACACATTCAGTTGGATGTTAAGAAAGAGAGGAATTGATGGGTTTTGAGGTTT 183
Db 121 TTTACACATTCAGTTGGATGTTAAGAAAGAGAGGAATTGATGGGTTTTGAGGTTT 180
OY 184 AAACCTTAAAGTACGACAGATTAAGTATGTTTACGTTGCTCTATATGTGTAAT 243
Db 181 AAACCTTAAAGTACGACAGATTAAGTATGTTTACGTTGCTCTATATGTGTAAT 240
OY 244 GAAGGCACTCCAAAGGTTCTTAGGTGATAGATTATTTAGACGATTTAATCATATA 303
Db 241 GAAGGCACTCCAAAGGTTCTTAGGTGATAGATTATTTAGACGATTTAATCATATA 300
OY 304 GTCCGTGGCGACTGTACATCATAGATTGTTTTTATTTTTCAGTAGCTGTGATGTT 363
Db 301 GTCCGTGGCGACTGTACATCATAGATTGTTTTTATTTTTCAGTAGCTGTGATGTT 360
OY 364 TTTTGATTTAATTACTATCTCAAAATCCAAATTCCTAAGACGACCAAG 423
Db 361 TTTTGATTTAATTACTATCTCAAAATCCAAATTCCTAAGACGACCAAG 420
OY 424 TCTCTCAATATGTAAACAGAACAAAGTTTTTGTAGTACCTAAAAGACACTCCCATG 483
Db 421 TCTCTCAATATGTAAACAGAACAAAGTTTTTGTAGTACCTAAAAGACACTCCCATG 480
OY 484 GAAGCTTGAGTCACTGTGGGATTTGGTCTCTCCCATTCCAATTATGCGATTACACAG 543
Db 481 GAAGCTTGAGTCACTGTGGGATTTGGTCTCTCCCATTCCAATTATGCGATTACACAG 540
OY 544 GCGAGCAAAAGCTCCGAGTAGCAACAACCTTACAACTATCTGCTCCGACGACAA 603
Db 541 GCGAGCAAAAGCTCCGAGTAGCAACAACCTTACAACTATCTGCTCCGACGACAA 600
OY 604 TGGGCGACGCGCTTCTCCGAGCTTCAATTACCTCGATTCCTCTCCCTCCCTTC 663
Db 601 TGGGCGACGCGCTTCTCTCCGAGCTTCAATTACCTCGATTCCTCTCCCTCCCTTC 660
```

```
OY 664 GCCACCGCCACCAACACCGGCACTGCTCTCTGACCATCTATGATCGTCCGAA 723
Db 661 GCCACCGCCACCAACACCGGCACTGCTCTCTGACCATCTATGATCGTCCGAA 720
OY 724 CGCCACGCTCCCATCCCATTTGATTTCTACCAAGTATTAAGACTGCAACAATTCCTTA 783
Db 721 CGCCACGCTCCCATCCCATTTGATTTCTACCAAGTATTAAGACTGCAACAATTCCTTA 780
OY 784 ACCGATGAATTCAGAAAGAGATTCGAAGCTAGGGTTTCGAAACCGCGCAATTCGTTTC 843
Db 781 ACCGATGAATTCAGAAAGAGATTCGAAGCTAGGGTTTCGAAACCGCGCAATTCGTTTC 840
OY 844 AGCAGCAACGCTTATACACCGGAGACAGATTTCTCAAGCTGCTCGAAACTGTCT 903
Db 841 AGCAGCAACGCTTATACACCGGAGACAGATTTCTCAAGCTGCTCGAAACTGTCT 900
OY 904 AATCCCTGGCTGTAGAAAGAGTCAAGAAAGTCTCTTATGATGAAGAAAGCTACAGTC 963
Db 901 AATCCCTGGCTGTAGAAAGAGTCAAGAAAGTCTCTTATGATGAAGAAAGCTACAGTC 960
OY 964 ATCACTGATGTCCTTGGGATTAAGTAAATTTTCGAAATTAATAAATTTCTTCGTT 1023
Db 961 ATCACTGATGTCCTTGGGATTAAGTAAATTTTCGAAATTAATAAATTTCTTCGTT 1020
OY 1024 TTAATTTCAATGATTTGATTAAGAGAGAACTTTTATCTAGTAAAGTTCCTGAGGCTCT 1083
Db 1021 TTAATTTCAATGATTTGATTAAGAGAGAACTTTTATCTAGTAAAGTTCCTGAGGCTCT 1080
OY 1084 CTGCTGATTCGAAAGAGTGTGAGACTGATAGTCTTCGAGTGTGAGAGCTGCTGCT 1143
Db 1081 CTGCTGATTCGAAAGAGTGTGAGACTGATAGTCTTCGAGTGTGAGAGCTGCTGCT 1140
OY 1144 TTAAGAGAGGTTGCTTAAGTCTGTTTAAAGAGATGTTGTTTAAAGTCTGCTGCT 1203
Db 1141 TTAAGAGAGGTTGCTTAAGTCTGTTTAAAGAGATGTTGTTTAAAGTCTGCTGCT 1200
OY 1204 TCTGATGTCCTCAAGGATGCTATGCGATGATTCACCTGATTTTATCTAGGTTATGA 1263
Db 1201 TCTGATGTCCTCAAGGATGCTATGCGATGATTCACCTGATTTTATCTAGGTTATGA 1260
OY 1264 GTTTGTTAGAGAGCTTTGAAGCTTTTACAGTAGTGTGCTGCTTGGTAAATTTGACG 1323
Db 1261 GTTTGTTAGAGAGCTTTGAAGCTTTTACAGTAGTGTGCTGCTTGGTAAATTTGACG 1320
OY 1324 AGCGTGGCTTTAAGAACTTTCTGATTTGATCTTGTATATGAGTCTTGTGAGGA 1383
Db 1321 AGCGTGGCTTTAAGAACTTTCTGATTTGATCTTGTATATGAGTCTTGTGAGGA 1380
OY 1384 GAAAGAGCAAGTAGCTTGCACCGGATTTACGTGCACAAATTTGATGAGACTTTGGAAGA 1443
Db 1381 GAAAGAGCAAGTAGCTTGCACCGGATTTACGTGCACAAATTTGATGAGACTTTGGAAGA 1440
OY 1444 GATGACTCCGCGTATGCTTTGAGAGTACTTGGCTTACCGCTTGAGATTAACGCTGC 1503
Db 1441 GATGACTCCGCGTATGCTTTGAGAGTACTTGGCTTACCGCTTGAGATTAACGCTGC 1500
OY 1504 GAAAAGACTAAATGTTTAAAGCGGTGCGGAAATATTTGTGCTGTTGAGAGAGTGG 1563
Db 1501 GAAAAGACTAAATGTTTAAAGCGGTGCGGAAATATTTGTGCTGTTGAGAGAGTGG 1560
OY 1564 AGCATCAGCTCTTGTGGGGGTTTGAACCGGTGAGAAATTTATGATGAGCGTTTTTACG 1623
Db 1561 AGCATCAGCTCTTGTGGGGGTTTGAACCGGTGAGAAATTTATGATGAGCGTTTTTACG 1620
OY 1624 AATGACAGCTGCTGAGAGATTAACAGTTTAACTTTTAAATTTCTTAAAGATGA 1683
Db 1621 AATGACAGCTGCTGAGAGATTAACAGTTTAACTTTTAAATTTCTTAAAGATGA 1680
OY 1684 TATPACTTTAGGTTTTCATTTTAAATGATGTTGTGAGTGGATGATCTTTTGTAGCT 1743
Db 1681 TATPACTTTAGGTTTTCATTTTAAATGATGTTGTGAGTGGATGATCTTTTGTAGCT 1740
OY 1744 ACCCAAGCAATATTCAGAGAGTCAATTTGAAGTTTACGAAGTTGCACTTGTCTTGTG 1803
```


Db 1741 ACCGAGCAATATTCAGAGATCTTGAAGTTACGAAGTTCCACTGCTCTG 1800
Qy 1804 GCTCAAGCTTTTATGTAAGAACCAACCTTTTACAGGATGCTGAAGCAATTCAG 1863
Db 1801 GCTCAAGCTTTTATGTAAGAACCAACCTTTTACAGGATGCTGAAGCAATTCAG 1860
Qy 1864 CAACCTTCAAGGCTAAGGTAATGCTATGAGAAATCTGCGATGTTGTAATGACGG 1923
Db 1861 CAACCTTCAAGGCTAAGGTAATGCTATGAGAAATCTGCGATGTTGTAATGACGG 1920
Qy 1924 AATTAATGGAGATTAACCTTCCGTCTAGAAAGGGGACTCTGACATGCTTATAGCAA 1983
Db 1921 AATTAATGGAGATTAACCTTCCGTCTAGAAAGGGGACTCTGACATGCTTATAGCAA 1980
Qy 1984 GTTATGAAATGCGATGATGTTGGGCTTGAAGAGGAGATTCACAAATATAGCAATCA 2043
Db 1981 GTTATGAAATGCGATGATGTTGGGCTTGAAGAGGAGATTCACAAATATAGCAATCA 2040
Qy 2044 GCTAATGTAAGTTGTTTGAAGAAATTCGAATCGTATGACAAATGATATCTCCCTGA 2103
Db 2041 GCTAATGTAAGTTGTTTGAAGAAATTCGAATCGTATGACAAATGATATCTCCCTGA 2100
Qy 2104 CTATGCAATTTGTTGAAACCTGCTGGGAGGAGGCTCTTCTTCTTGGTTCAAGACACC 2163
Db 2101 CTATGCAATTTGTTGAAACCTGCTGGGAGGAGGCTCTTCTTCTTGGTTCAAGACACC 2160
Qy 2164 AAGATTAATAATTTTAACTCGGAGGACTATGATGATCTATGTTTGAATCTG 2223
Db 2161 AAGATTAATAATTTTAACTCGGAGGACTATGATGATCTATGTTTGAATCTG 2220
Qy 2224 GAAAGATGAGGATGTTGAGGTTCTCTTTAGCTGCTGCTGCACTATGCAAGGAT 2283
Db 2221 GAAAGATGAGGATGTTGAGGTTCTCTTTAGCTGCTGCTGCACTATGCAAGGAT 2280
Qy 2284 GGAAGCGACATGTAAGAGCTATGCTATGCAAGGCACTGCAAGAAATTTTCTCCGC 2343
Db 2281 GGAAGCGACATGTAAGAGCTATGCTATGCAAGGCACTGCAAGAAATTTTCTCCGC 2340
Qy 2344 TATACAGATAGAACTCGGCTGAACCCAGAGATGTCAGAGAGAGCTTTAGTATGAT 2403
Db 2341 TATACAGATAGAACTCGGCTGAACCCAGAGATGTCAGAGAGAGCTTTAGTATGAT 2400
Qy 2404 CCTGTTGTAACATGTAAGCCGCTGATGTAAGCCGTGCTCTTAATTCAGAGCTGA 2463
Db 2401 CCTGTTGTAACATGTAAGCCGCTGATGTAAGCCGTGCTCTTAATTCAGAGCTGA 2460
Qy 2464 AAGCCCTCTGAAGCTTTGAACTATGATTTGCAATTCGAGCTGGGCTCTCAGAGAT 2523
Db 2461 AAGCCCTCTGAAGCTTTGAACTATGATTTGCAATTCGAGCTGGGCTCTCAGAGAT 2520
Qy 2524 AAGCGTTGATGAATCTGCTGTAATGTCGTTGCTGATATGTTAAAGGAGGCAAGTGG 2583
Db 2521 AAGCGTTGATGAATCTGCTGTAATGTCGTTGCTGATATGTTAAAGGAGGCAAGTGG 2580
Qy 2584 AAGATCTTAGCTGCTGCTGTAATGTAAGCTGATTCCTGCTGCTGCTGCTGCTGCT 2643
Db 2581 AAGATCTTAGCTGCTGCTGTAATGTAAGCTGATTCCTGCTGCTGCTGCTGCTGCT 2640
Qy 2644 CTTTAAAGAGCTCATCTTTTCAAGCAAGATATGCTTCTTATGAGATCTGATGTC 2703
Db 2641 CTTTAAAGAGCTCATCTTTTCAAGCAAGATATGCTTCTTATGAGATCTGATGTC 2700
Qy 2704 GCTACCATGATGATTAATGATGCAATTTTCAATATTCGCAATGCTCAAAATATGC 2763
Db 2701 GCTACCATGATGATTAATGATGCAATTTTCAATATTCGCAATGCTCAAAATATGC 2760
Qy 2764 TTGTTTGTAGCTAAGACATATGTTCCACTTAATATACATGCTCCAAAGTTGTAACAG 2823
Db 2761 TTGTTTGTAGCTAAGACATATGTTCCACTTAATATACATGCTCCAAAGTTGTAACAG 2820
Qy 2824 ATTAACAAGTTGCTGATTAATTTCACTAATATGCTGCTGATTTTGTCAACTG 2883

Db 2821 ATTAACAAGTTGCTGATTAATTTCACTAATATGCTGCTGATTTTGTCAACTG 2880
Qy 2884 TAGACAGAAATGTAATTTCACTCAACATTTCTGTTAGAAATACGATGATAGAGA 2943
Db 2881 TAGACAGAAATGTAATTTCACTCAACATTTCTGTTAGAAATACGATGATAGAGA 2940
Qy 2944 TTGCTTATGCTGCTTTGCTCAACTTTTCTTCTTGAATTTTCTTTGATTTAG 3003
Db 2941 TTGCTTATGCTGCTTTGCTCAACTTTTCTTCTTGAATTTTCTTTGATTTAG 3000
Qy 3004 GTTACGACAGCTGACGATGCAATGCAAGCACTCCCAATATGATGCTAGAGACAGAGA 3063
Db 3001 GTTACGACAGCTGACGATGCAAGCACTCCCAATATGATGCTAGAGACAGAGA 3060
Qy 3064 TATATGATTCAGATGAGCAAGATTAATGCTCTGCTTTTGGGCTGATACCGCATAGA 3123
Db 3061 TATATGATTCAGATGAGCAAGATTAATGCTCTGCTTTTGGGCTGATACCGCATAGA 3120
Qy 3124 AATGTTACAGAGGTAAGGAAATTAATCTCAATTCATATGTTGTAATGTTGG 3183
Db 3121 AATGTTACAGAGGTAAGGAAATTAATCTCAATTCATATGTTGTAATGTTGG 3180
Qy 3184 ACATGATTAATGCTGCTGCTGCTGTTGATTTCTGTTATTTATGATTTGATGAGCCGA 3243
Db 3181 ACATGATTAATGCTGCTGCTGCTGTTGATTTCTGTTATTTATGATTTGATGAGCCGA 3240
Qy 3244 TGCTAAGATTTGACATGACAGAGGCTGAACTGCGAGCTTGAGTTGATTTGAT 3303
Db 3241 TGCTAAGATTTGACATGACAGAGGCTGAACTGCGAGCTTGAGTTGATTTGAT 3300
Qy 3304 ATACATGTTGAACTATCTGTTGACAGTGTGACAGTCTGACAGATGAAACCGTGTCTC 3363
Db 3301 ATACATGTTGAACTATCTGTTGACAGTGTGACAGTCTGACAGATGAAACCGTGTCTC 3360
Qy 3364 TGTGGAAGCAACTCTGGAAGAGTCTGTTGCTATCTGATTTGATTTGATTTGATTTGAT 3423
Db 3361 TGTGGAAGCAACTCTGGAAGAGTCTGTTGCTATCTGATTTGATTTGATTTGATTTGAT 3420
Qy 3424 ATGCTACTGATGTCGAGCACTACACAAAGATGAGATTTTCTGCTCAAGTCAAGGCT 3483
Db 3421 ATGCTACTGATGTCGAGCACTACACAAAGATGAGATTTTCTGCTCAAGTCAAGGCT 3480
Qy 3484 GGAATATCACTGAAGGCTCTGTTCTTGCAATTAATATGATATGATGATGATGATGAT 3543
Db 3481 GGAATATCACTGAAGGCTCTGTTCTTGCAATTAATATGATATGATGATGATGATGAT 3540
Qy 3544 CTGCGAATTTCTGTTGCTGTAATTTCTCTTAAGTTAGTTTATTAATGAACA 3603
Db 3541 CTGCGAATTTCTGTTGCTGTAATTTCTCTTAAGTTAGTTTATTAATGAACA 3600
Qy 3604 CAAAAAATTAACGTTCTTGGAACACCTTTCTGATCTAATCTAATACATAAGGCT 3663
Db 3601 CAAAAAATTAACGTTCTTGGAACACCTTTCTGATCTAATCTAATACATAAGGCT 3660
Qy 3664 ACA 3667
Db 3661 ACA 3664

RESULT 2
AY221467
LOCUS
DEFINITION Arabidopsis thaliana truncated division protein (arce6) gene, arce6-1
allele, complete cds; nuclear gene for chloroplast product.
ACCESSION
VERSION AY221467.1 GI:3436274
KEYWORDS
SOURCE Arabidopsis thaliana (chale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE
1 (bases 1 to 3664)

QY 1504 GAAAAGACTTAATGCTTTAAGCGGTGCGGAATATTTTGTGTCCTTTGAGAGAGTGG 1563
DB 1501 GAAAAGACTTAATGCTTTAAGCGGTGCGGAATATTTTGTGTCCTTTGAGAGAGTGG 1560
QY 1564 AGCATAGCTCTTGTGGGGGTTTGACCGGTGGAAGTTTATGAANTGAGCGCTTTTAG 1623
DB 1561 AGCATAGCTCTTGTGGGGGTTTGACCGGTGGAAGTTTATGAANTGAGCGCTTTTAG 1620
QY 1624 AATGACAGCTGCGAGCAGGTATACAGTTTAACTCTTTTAACTTTTAACTTTTAACTAG 1683
DB 1621 AATGACAGCTGCGAGCAGGTATACAGTTTAACTCTTTTAACTTTTAACTTTTAACTAG 1680
QY 1684 TATTAATCTTAAGTTTCTCATTTTAAATGTAATGTTGTGCTGAGTTGATCTTTTGTAGCT 1743
DB 1681 TATTAATCTTAAGTTTCTCATTTTAAATGTAATGTTGTGCTGAGTTGATCTTTTGTAGCT 1740
QY 1744 ACCCCAGCAATATTCACAGAGTCACTTGAAGTTTGAAGTTGACCTTGCTGTGG 1803
DB 1741 ACCCCAGCAATATTCACAGAGTCACTTGAAGTTTGAAGTTGACCTTGCTGTGG 1800
QY 1804 GCTCAAGCTTTTATTTGTAAGAACCACTTTTACAGAGTCTGATAGCAATTCAG 1863
DB 1801 GCTCAAGCTTTTATTTGTAAGAACCACTTTTACAGAGTCTGATAGCAATTCAG 1860
QY 1864 CAACCTTCAAGCAGGCTAAGGTAAATGCTATGAGATTCCTGCAATGTTTATGATACAGG 1923
DB 1861 CAACCTTCAAGCAGGCTAAGGTAAATGCTATGAGATTCCTGCAATGTTTATGATACAGG 1920
QY 1924 AATTAATGGGAGATGACCTTGCTGTAAGAAAGGGGACTGTGCACTGCTTATAGGCAAA 1983
DB 1921 AATTAATGGGAGATGACCTTGCTGTAAGAAAGGGGACTGTGCACTGCTTATAGGCAAA 1980
QY 1984 GTTGAATGATGCGGATGATGTTGGGCTTGAACAGTGAAGATTCATAATAGCAATCCA 2043
DB 1981 GTTGAATGATGCGGATGATGTTGGGCTTGAACAGTGAAGATTCATAATAGCAATCCA 2040
QY 2044 GCTATTTGGAGTTGTTTGAAGAAATTCGAATCGTATGACAAATGATGATCTCCCTGGA 2103
DB 2041 GCTATTTGGAGTTGTTTGAAGAAATTCGAATCGTATGACAAATGATGATCTCCCTGGA 2100
QY 2104 CTATGCAATTTGTTGAAAACCTGTTGCGAGGGGTTGCTTCCATAGTTGCAAGAACCC 2163
DB 2101 CTATGCAATTTGTTGAAAACCTGTTGCGAGGGGTTGCTTCCATAGTTGCAAGAACCC 2160
QY 2164 AAAAGTAAAAAATTTAACTCGGGGACTACTATGATGATCTATGATTTTGAATCTTG 2223
DB 2161 AAAAGTAAAAAATTTAACTCGGGGACTACTATGATGATCTATGATTTTGAATCTTG 2220
QY 2224 GAAAGATGAGGTAGTTCAAGGTTCTCCTTTAGCTGCTGCAACTATGCGAAGAT 2283
DB 2221 GAAAGATGAGGTAGTTCAAGGTTCTCCTTTAGCTGCTGCTGCAAGTATGCGAAGAT 2280
QY 2284 GGAAGCGAGCATGTTGAAAAGTATGCTATGACGACCTGCGAAGATTTTCTTCCCGC 2343
DB 2281 GGAAGCGAGCATGTTGAAAAGTATGCTATGACGACCTGCGAAGATTTTCTTCCCGC 2340
QY 2344 TATACAGATAGAACTCGGCTGAACCCAAAGATGTCAGAGACAGCTGTTTATGTTAGAT 2403
DB 2341 TATACAGATAGAACTCGGCTGAACCCAAAGATGTCAGAGACAGCTGTTTATGTTAGAT 2400
QY 2404 CCTGTTGTAACATGTAAGCCGTATGAGTGAAGCTGCTGCTTTTATTTGACAGACTGTA 2463
DB 2401 CCTGTTGTAACATGTAAGCCGTATGAGTGAAGCTGCTGCTGCTTTTATTTGACAGACTGTA 2460
QY 2464 AGACCTCTGAAAACCTTTGAAACTAATGATTAATGCAATTCAGCTGGGCTCTCAAGAGT 2523
DB 2461 AGACCTCTGAAAACCTTTGAAACTAATGATTAATGCAATTCAGCTGGGCTCTCAAGAGT 2520
QY 2524 AGCGTTGATGAATACTACTGTTGAATGTCGTTGCTGATATGTTAAAGAGGCAAGTGTG 2583
DB 2521 AGCGTTGATGAATACTACTGTTGAATGTCGTTGCTGATATGTTAAAGAGGCAAGTGTG 2580
QY 2584 AAGATCTAGCTGCTGCTGCTGCAATTGCACTGTTCAAGCCAGAAATATTTT 2643

DB 2581 AAGATCTAGCTGCTGCTGCTGCAATTGCACTGATTTCACTGTTCAAGCCAGAAATATTTT 2640
QY 2644 CTTAAGACAGCTCACTCTTTTCAAGCAGATATGTTTCTTCTATGAAATCTGATGTC 2703
DB 2641 CTTAAGACAGCTCACTCTTTTCAAGCAGATATGTTTCTTCTATGAAATCTGATGTC 2700
QY 2704 GCTACATAGGTATGATTAATTAATGATGCAATTTGATATATCTGATCTCAAAATATGTC 2763
DB 2701 GCTACATAGGTATGATTAATTAATGATGCAATTTTCAATATATCTGATGCTCAAAATATGTC 2760
QY 2764 TTTGTTTGTGAGCTTAAGAACTATGTTCCACTTAATATACATGTCGCAAAAGTTGTAACAG 2823
DB 2761 TTTGTTTGTGAGCTTAAGAACTATGTTCCACTTAATATACATGTCGCAAAAGTTGTAACAG 2820
QY 2824 AATTAACAAGTTGCTGAGTAATTTCACTTAATATGCTGCTGAATTTTGTATCAAACTG 2883
DB 2821 AATTAACAAGTTGCTGAGTAATTTCACTTAATATGCTGCTGAATTTTGTATCAAACTG 2880
QY 2884 TAGACAGAAATGTAATTTCACTCTCAACATTTCTGTTTGAATTAAGTATGATTAAGAGA 2943
DB 2881 TAGACAGAAATGTAATTTCACTCTCAACATTTCTGTTTGAATTAAGTATGATTAAGAGA 2940
QY 2944 TTTGCTTATGTTGCTTTGTCCACTTTTCTTCTGATTTTCTTTTGTGATTTAGG 3003
DB 2941 TTTGCTTATGTTGCTTTGTCCACTTTTCTTCTGATTTTCTTTTGTGATTTAGG 3000
QY 3004 GTCACTCAGAGCTGACATTTGAGAAAGCCTCCAGAAATGATGCTGAGACTGACAGAGA 3063
DB 3001 GTCACTCAGAGCTGACATTTGAGAAAGCCTCCAGAAATGATGCTGAGACTGACAGAGA 3060
QY 3064 TATAGTATCCAAAGTGGCAGAAAGATTAAGTCTCTGGCTTTTGGGCTGATCCAGATAGA 3123
DB 3061 TATAGTATCCAAAGTGGCAGAAAGATTAAGTCTCTGGCTTTTGGGCTGATCCAGATAGA 3120
QY 3124 AATGTTACCAAGGTGAGGGAATTAATCTACATTTCAATCAATGTTGTGAAAACCTGTTGG 3183
DB 3121 AATGTTACCAAGGTGAGGGAATTAATCTACATTTCAATCAATGTTGTGAAAACCTGTTGG 3180
QY 3184 ACATGATTTATGTCGTTGGCTTGTGTAATCTGTTTATTAATAGTTTGGATGGGCAAA 3243
DB 3181 ACATGATTTATGTCGTTGGCTTGTGTAATCTGTTTATTAATAGTTTGGATGGGCAAA 3240
QY 3244 TGTGAAAGATTTGACCTGACAGACAGACTGAACTGCGCAGCTTGAGTTTATGATTT 3303
DB 3241 TGTGAAAGATTTGACCTGACAGACAGACTGAACTGCGCAGCTTGAGTTTATGATTT 3300
QY 3304 ATPACACTGTTGAAAATCTATCTGTTGACAGTGTGACAGTCTCAGCAGATGAAACCCGTGCTC 3363
DB 3301 ATPACACTGTTGAAAATCTATCTGTTGACAGTGTGACAGTCTCAGCAGATGAAACCCGTGCTC 3360
QY 3364 TGTGTAAGAAAGCACTCTGAGAGAGTCTGTTGCTATCTGATTTGTTGTTTCAAGAAA 3423
DB 3361 TGTGTAAGAAAGCACTCTGAGAGAGTCTGTTGCTATCTGATTTGTTGTTTCAAGAAA 3420
QY 3424 ATGCTACTGATGTCAGAACTTACACAACAAGATGCAAGTTTCTGTTCCAAAGTCAAGGT 3483
DB 3421 ATGCTACTGATGTCAGAACTTACACAACAAGATGCAAGTTTCTGTTCCAAAGTCAAGGT 3480
QY 3484 GGAATACTGTAAGGCTCTGTTCTTGCACTCAATATTAATCTCATATGATGATGCTGAG 3543
DB 3481 GGAATACTGTAAGGCTCTGTTCTTGCACTCAATATTAATCTCATATGATGATGCTGAG 3540
QY 3544 CTTGCGAGATTTCTTGTGTTCTGTAATTTCTCTCAAGTTTATGTTTAAATGTAACA 3603
DB 3541 CTTGCGAGATTTCTTGTGTTCTGTAATTTCTCTCAAGTTTATGTTTAAATGTAACA 3600
QY 3604 CAAAAAATTAATCACTTCTTGACACACCTTTTCTTATCTAATCAATTAAGGCT 3663
DB 3601 CAAAAAATTAATCACTTCTTGACACACCTTTTCTTATCTAATCAATTAAGGCT 3660
QY 3664 ACAA 3667
|||||

Db 3661 ACAA 3664

RESULT 3
LOCUS AB016888
DEFINITION Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:MDH9.
ACCESSION AB016888 BA000015
VERSION AB016888.1 GI:3449329
KEYWORDS
SOURCE
ORGANISM Arabidopsis thaliana (thale cress)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1
Asanizu,E., Sato,S., Kaneko,T., Nakamura,Y., Kotani,H., Miyajima,N. and Tabata,S.
Structural analysis of Arabidopsis thaliana chromosome 5. VIII. Sequence features of the regions of 1,081,958 bp covered by seventeen physically assigned P1 and TAC clones
DNA Res. 5 (6), 379-391 (1998)
10048488
2 (bases 1 to 85791)
Nakamura,Y.
Direct Submission
Submitted (18-AUG-1998) Yasukazu Nakamura, Kazusa DNA Research Institute, Department of Plant Gene Research; 1532-3, Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:ynakamu@kazusa.or.jp, tel:81-438-52-3935, Fax:81-438-52-3934)
Address for correspondence: kaos@kazusa.or.jp
For the latest information on annotation of this clone, please see http://www.kazusa.or.jp/kaos/cgi-bin/agd_graph.cgi?c=MDH9
Genes with similarity to proteins in the databases are described in 'product' or 'note' qualifiers. Genes that have no significant protein similarity are described as 'unknown protein'.
The software programs used to predict genes include: Grail (Informatics Group, Oak Ridge National Laboratory, <http://compbio.ornl.gov/Grail-1.3/>), GENSCAN (Chris Burge, MIT, <http://CCR-081.mit.edu/GENSCAN.html>), NeGene2 (S.M. Hebsgaard, et al., CBS, Technical University of Denmark, <http://www.cbs.dtu.dk/services/NeGene2/>) and SplicePredictor (Volker Brendel, Stanford University, <http://grem1ml.zool.iastate.edu/cgi-bin/sp.cgi>).
Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, Washington University School of Medicine, St. Louis, <http://genome.wustl.edu/eddy/tRNAscan-SE/>).
This sequence may not be the entire insert of this clone. It may be shorter because we remove overlaps between neighboring submissions.
The 5' clone is K5J14 and the 3' clone is K16E1.
Location/Qualifiers
1. 85791
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/chromosome="5"
/clone="MDH9"
/clone_1fb="Witeui P1"
/ecotype="Columbia"
/complement(59. .239)
/note="CDS is reported in Acc# AB023032
gene_id:K5J14.12
unknown protein"
/number=1
/evidence=not experimental
complement(join(2056. .2079,2278. .2393,2991. .3633))
/note="contains similarity to unknown protein
gene_id:MDH9.1
pir||T04864"
/pseudo
/codon_start=1
/evidence=not experimental
join(5739. .6138,6462. .6595)

CDS

/note="unnamed protein product; gene_id:MDH9.2
unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAB10474.1"
/db_xref="GI:9759469"
/translation="MKDEPAEEEDKSSNLSESLAKNSVPLPERGRFYEEYAKNE
RLRRKGEEAVEGGVKGTPTNLGVNHEPTNRRGTAKESIKKTVSVSESSIPR
YSLSMAKENRPPPLPLNAVVSAMKSVTYRKEERLLCTCACTSSGLAFLPRLTGS
EPLGVHKKRTLPFSALV"
complement(join(6868. .7911,7989. .8414,8555. .8671,
8754. .9137))
/note="gene_id:MDH9.3"
/codon_start=1
/evidence=not experimental
/product="varm repeat containing protein"
/protein_id="BAB10475.1"
/db_xref="GI:9759470"
/translation="WDTDEATGDAENRDEVTAEERPIHDEVDAVEIHEEYKEDD
DCGLVSDIVSIFLDQINGRYRTOQCEFNIVRLKILIPLEDIRGESPSCNF
LNLRLKVFLLAKKLLLETCSNGSKTYVALDETMTFRSHIYEKLNVLVAPDEKMI
SGAKOEDSLCKOLKKAKRRTDODIELADVMVVSFKDPPNADSAIIERLAKLE
LOTIDLKETTTAIOSLIODKGLNIETKOHIELLNKFKLQLEATDII.YOPIVK
AITKSTSLIIPHEFLCPTITELMDPIYITAGQTYEKESQKQFADGHTCPTROEL
DHLSLAPNFAKLNLMQWCEKNFKIPEKVSPOSOEQDEVSLVEALSSQLEQ
RSYVKQKRLARENENRVLIVANAGAIPLVQLSLSPDSGIQNAVYTLNLISIDEV
KSLISNGALPNIIEIENGREARNSAALFSLMLDENKVTIGISNGIPLVDIL
OGLTRKSCALFALFSLNSANKRAIDAGIYQPLNLNKKNGIMIDEALSIIL
LASHPEROALIGOLSFETLVEFTIGRTPNKCAVSLLEKSNSSFIILAQGV
YEVIVETTSQRAQKNAALIGLSKSGT"
10888. .12579
/note="unnamed protein product; gb|AAF07802.1
gene_id:MDH9.4
similar to unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAB10476.1"
/db_xref="GI:9759471"
/translation="MMISEKPLGEBSIRDLLEVLTVSRRLVKSVOQLKKIKHTEV
VEBEIRAGAVNCLISVGCRAVDGDEPDSSNKSWSASBEGKGLMTICGEERL
DCPSYGRERFMKNNRKYLADSGODRKHYVLPDILMCLRLPLTSLNLNLAK
KMSMANTQFLMRREGSFQTPMLFPAALXGSCSGDHYGVDSODKKHRIETDL
KGRFMSVTSIHIEIYIVGSRMDRNSFKSHRGLIFSPSITAKRKTIASRRHRSAPI
VGAITEVSEPTVQTKONRODRPHLSRVGSESDVLEDPRLSVRONRNSADONGTK
SHRLIRKQRLNRNSKRFVLAIGGTGFDEPLDSGETYDSATNTWSMOELPMGF
GVVSCGIIICNGIFYAYSENDKLSGYDIERGFMTITQSPIPRVHEFTYPLVSCNHL
FMLSVCSDGSDGIGBRNKAVKRLMELDLVLTWTEVSHPPADPMNATVYSDQVI
LMGIEWKIRGVLSPTVCDVLTTEASRHWNRNQRSLNLSCTNKITALLHL"
complement(15087. .16778)
/note="unnamed protein product; gb|AAF07802.1
gene_id:MDH9.5
similar to unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAB10477.1"
/db_xref="GI:9759472"
/translation="MMISEKPLGEBSIRDLLEVLTVSRRLVKSVOQLKKIKHTEV
EDEEVARGAVNCLISVGCRAVDGDEPDSSNKSWSASBEGKGLMTICGEERL
DCPSYGRERFMKNNRKYLADSGODRKHYVLPDILMCLRLPLTSLNLNLAK
KMSMANTQFLMRREGSFQTPMLFPAALXGSCSGDHYGVDSODKKHRIETDL
KGRFMSVTSIHIEIYIVGSRMDRNSFKSHRGLIFSPSITAKRKTIASRRHRSAPI
VGAITEVSEPTVQTKONRODRPHLSRVGSESDVLEDPRLSVRONRNSADONGTK
SHRLIRKQRLNRNSKRFVLAIGGTGFDEPLDSGETYDSATNTWSMOELPMGF
GVVSCGIIICNGIFYAYSENDKLSGYDIERGFMTITQSPIPRVHEFTYPLVSCNHL
FMLSVCSDGSDGIGBRNKAVKRLMELDLVLTWTEVSHPPADPMNATVYSDQVI
LMGIEWKIRGVLSPTVCDVLTTEASRHWNRNQRSLNLSCTNKITALLHL"
complement(join(18111. .18193,18281. .18535,18773. .19001,
19088. .19195,19307. .19447,19542. .19952,20129. .20245))
/note="unnamed protein product; gene_id:MDH9.6
similar to unknown protein
bp|P42251"
/codon_start=1
/evidence=not experimental

/protein_id="BAB10478.1"
/db_xref="GI:9759473"
/translation="MNQISFMCISLLITVIAAEETPTIRIAFGSCANQSAPOPIWDA
INKFPOLFIWGDNIYGDRIIRPLVIGERTFGPRNRPPLVPSNEEMKRYAKAK
ANPYSRIORAKVIGMDHDHGLNDACKERDRKINOKMLDPFNEELDSRRKOA
GVSYVTVGPNRKVVVLDTRHYRDLPSGSLIDPDMWIMNEISGPSSETII
GSSVQVYISLSLSTTGPFLYMESWGRPPKRRKFLKIDATKRDGVIPLISGDVHPEIT
RYDCSVGYPDYVTSGLVQSVKVPRLRSIIRLLFYTSTYKRVINDNCKFKSCT
YQONFGAISIDMANPVTIKLEIRVDNDIVLGTWLSSEIQPGSNLKTOTTKGK
SQRCTLEIELEPTTRVLAVLIYFTIATAMAILGLIIGAVIATACVYCKKVD"
complement (21540..22097)
/note="unnamed protein product; contains similarity to
calmodulin
gene_id:MDH9.7"
/codon_start=1
/evidence=not experimental
/protein_id="BAB10479.1"
/db_xref="GI:9759474"
/translation="MTLAKNQSLSRLYKYSKRSESRNLEDESRTSNSGSSSS
LAVNELRTVFDYDANSDGKISGEELQCVSLGGALSGREVEVVKTSVDVGDGFID
FEEFLKMEGEDSDERRKELKEAFGMVMEGEERTTAASIRRTLRLSGESCVDAC
KMIKGFNDNDGVLSDFEFLMMR"
join(24090..24289,24410..24614,24701..24945,25115..25313,
25632..25679,25821..25928,26039..26227,26314..26613,
26730..26831,27169..27303,27487..27711,27919..28122,
28213..28380,28646..28795,28889..29026,29204..29431,
29519..29671,29784..29910,30115..30241,30554..30674,
30763..30843,30932..31033,31112..31262,31338..31429)
/note="gene_id:MDH9.8"
/codon_start=1
/evidence=not experimental
/product="pili1lysin"
/protein_id="BAB10480.1"
/db_xref="GI:9759475"
/translation="MASSSSSTFVGKSPILAPNSGDSRPRVYAKDSRNVKPNPS
SRRLPHRYVAPSLIIPYNGIMAAQPNSHKRLKRNYSGEATGISLSQRNCLT
CRRNAGIRRALPSAFVDRTAFLSRSSLTSLRHSQIVNATLIDPEHAGATMPD
GIVARODLDLIPRISDALEAFLEAGLPSHPKLRQKNGKLYLILPNVFNPRF
EAMHVVHVSIDEEDEQGIAMIEHVAFLPSKREKLGITGARANAAATDPFHVTHI
HSPITKXSEBDLPSPVLDALNEIAHNPFLSSRVKERRALISELMMNTYSEVYDC
OLLOHSENKIGRRPPIGLEROIKKMDVKTRKHEHMYPPANNTIIVYGDINDPR
LYHNIEAVGKGLDNESSTPSPSPGAPMANPLVPKIPAGLGTFPSEKNTINDOS
KMKIKERHAIKRPVHNMSLPGTSVLDLPQIFKHLLQNFALINFCIKIPKVTGFG
DLRNLMKRIFLSAHFRINTRYKSNBPPEVLELHDSGSECTVTTLITVAPON
MGNAYVAVQEVRLKEPVTGELTRVMDALKDSEHIAAMIDVSVNDIDFMS
DALSHVMDQOHEHLEAVAGTVLLEEVNAGVLESTISPGTAPLPALVAVY
PKVYHVGESDPMISPEILIESVKSLGAIPEAEPELVEKELISOSOLKEFLOR
NCPYPIRSGSLTKLHDKETGITQRLSGLIAVNYKKTSTERRAEMRLIVGGGAAR
TSDSGAVVGVRLTBSBGVDFSRBOVELFCVNLINCSLESTEEFLAMEFRFLR
DNGMAAFQRLMHVLERISVLEDAFDRAQLYSFRSIPKLEBATHKMLAMING
DERFVEPTPKSLQSLNLESVDAMVMSHFVGNMEVSIYVDFSEBIEHCIDLDGVK
ASHDSAKPPGSEPILEPROPTAGLOOQVFLKOTDRACGYIIPGAPLNNMGVDDDL
FOVSVKIPVAHGDILKSEBOLLGGDRLOKTRAPLFCVGTMLLAIEINSRLEPT
VRDSGLTYDVSPELNLPRNLGMYIVSYSTPKGYKADVACSVRGHNSDIAP
RELDAKRTLLMRHAELEKSNAYWLNLAHLQASSVPKRELSICEIVLSLEAASIED
IYLANOLRVDESDLYSCIGIAGAOAGEITVLSSEEBEDVPSICVLEVSGSSWTR
PTT"
complement (join(33251..33326,33512..33663,33744..33977,
33888..33968,34143..34265,34359..34768,34873..35052,
CDS
Query Match 99.8%; Score 3659; DB 15; Length 85791;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3662; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 TCTTCGCAATTAAGAGAAATACATTAATTAAGCAATTTGCTTGATTTCAACAAGATTTTG 60
DB TCTTCGCAATTAAGAGAAATACATTAATTAAGCAATTTGCTTGATTTCAACAAGATTTTG 63656
QY 61 CTTCGCTAATAGATTCATTCGCTGCTGTTGCTTTTACATTTACATGCTCATATAGTTTCG 120
DB CTTCGCTAATAGATTCATTCGCTGCTGTTGCTTTTACATTTACATGCTCATATAGTTTCG 63716
QY 121 AATTTTACATTTTCAATTTGATTTAGAAAAAGAGAGAAATTAATGATGGGCTTTTGTGGG 180

DB 63717 AATTTTACATTTTCAATTTGATTTAGAAAAAGAGAGAAATTAATGATGGGCTTTTGTGGG 63776
QY 181 TTTAACTTTAAAGTAGTCAAGAAATTAAGTCAATGTTTACTGTTGCTCTATATAGTAA 240
DB TTTAACTTTAAAGTAGTCAAGAAATTAAGTCAATGTTTACTGTTGCTCTATATAGTAA 63836
QY 241 AATGAAGCAACTCCACGGTTCTTAAGTGAAATGATATTTAGACGATTTAACATCAT 300
DB AATGAAGCAACTCCACGGTTCTTAAGTGAAATGATATTTAGACGATTTAACATCAT 63896
QY 301 AAAGTCGTGGCGACTGTAAATCATCATATGTTTTTTTATTTTTCAGTAGCTGTGAT 360
DB AAAGTCGTGGCGACTGTAAATCATCATATGTTTTTTTATTTTTCAGTAGCTGTGAT 63956
QY 361 GTTTTGTATTTAACTTATATCTATCAAAATCAAAATTCATTAACCTTAGACGACAAA 420
DB GTTTTGTATTTAACTTATATCTATCAAAATCAAAATTCATTAACCTTAGACGACAAA 64016
QY 63957 GTTTTGTATTTAACTTATATCTATCAAAATCAAAATTCATTAACCTTAGACGACAAA 64016
DB 64017 CAGTCTCTTCAATATGTAAAAAGAAAGATTTTGTAGTAGCCTTAAAAAGACCTCC 64076
QY 481 ATGGAAGCTCTGAGTCAAGTGGGCAATGCTCTCCCATTTCCATTTATGCCGATTACCA 540
DB ATGGAAGCTCTGAGTCAAGTGGGCAATGCTCTCCCATTTCCATTTATGCCGATTACCA 64136
QY 541 CCGGAGACGAAGAAAGCTCCGACGTAGGACCAACACCTTACAACTATGTGCTCCGACGC 600
DB CCGGAGACGAAGAAAGCTCCGACGTAGGACCAACACCTTACAACTATGTGCTCCGACGC 64196
QY 601 AAATGGCCGACCGCTCTCTCCGACCTTCAATTTCACTCGATTCCTCTCTCTCTCC 660
DB AAATGGCCGACCGCTCTCTCCGACCTTCAATTTCACTCGATTCCTCTCTCTCTCTCC 64256
QY 64197 AAATGGCCGACCGCTCTCTCCGACCTTCAATTTCACTCGATTCCTCTCTCTCTCC 64256
DB 661 TTCCGCAACCGCACCAACCGGCACTCTGCTCTCTGCGCAACATCTATGATGTCTCC 720
QY 64257 TTCCGCAACCGCACCAACCGGCACTCTGCTCTCTGCGCAACATCTATGATGTCTCC 64316
DB 721 GAAGCCGACGTCCTCCATCCCATTTGATTTTCAACAGGATTAAGGACCTCAACACATTTG 780
QY 64317 GAAGCCGACGTCCTCCATCCCATTTGATTTTCAACAGGATTAAGGACCTCAACACATTTG 64376
DB 781 TTAACCGATGAAATCAGAAAGCAATTCGAAAGTAAAGGTTTTCGAAATTCGAT 840
QY 64377 TTAACCGATGAAATCAGAAAGCAATTCGAAAGTAAAGGTTTTCGAAATTCGAT 64436
DB 841 TTACCGACGACGCTTTATATGACCGGACAGATTTTCAAGCTCTTGCGAAACTGTG 900
QY 64437 TTACCGACGACGCTTTATATGACCGGACAGATTTTCAAGCTCTTGCGAAACTGTG 64496
DB 901 TCTATCTCGGCTAGAAAGAGATGACAAAGGCTCTCTTGATGAGAAAGAGCTACA 960
QY 64497 TCTATCTCGGCTAGAAAGAGATGACAAAGGCTCTCTTGATGAGAAAGAGCTACA 64556
DB 961 GTCAATCACTGATGTTCTTGAGATTAAGGATTAATTCGATTCGAAATTAAGGTTTCTTC 1020
QY 64557 GTCAATCACTGATGTTCTTGAGATTAAGGATTAATTCGATTCGAAATTAAGGTTTCTTC 64616
DB 64617 GTTTTAAATTCATGAATTCGATTAAGGAGAACTTTATATCAGTGAAGGTTTCTCGGTGC 64676
QY 1081 TCTCGGTATTTGAAGAAAGGTGTGAGACTGAGATGTTCTTGCGGGTGTGAGGCTCT 1140
DB TCTCGGTATTTGAAGAAAGGTGTGAGACTGAGATGTTCTTGCGGGTGTGAGGCTCT 64736
QY 1141 GCTTAAGGAGAGTGTGCTTAAGTGTGTTTAAAGCAAGATGTGTTTAAATAGCGCTTGC 1200
DB GCTTAAGGAGAGTGTGCTTAAGTGTGTTTAAAGCAAGATGTGTTTAAATAGCGCTTGC 64796
QY 1201 GTTTCGATGTTCTGAGGAGATGCTATGAGCATTTGATTCACCTGATTTTATCTGCTTA 1260

Dh 64797 GTTCTCGATGTCGAGGAGTCATGCGATTGATCCACCTGATTTTATTAACGTGTTA 64856
Oy 1261 TGAAGTTGTGAGAAAGCTTTGAAAGCTTTTACAGGTAGTTGATGACTGCTTGGTAATTG 1320
Db 64857 TGAAGTTGTGAGAAAGCTTTGAAAGCTTTTACAGGTAGTTGATGACTGCTTGGTAATTG 64916
Oy 1321 ACGAGCGTTGGCTTTATAGAACCTTCTGTGAATTGATCTTGTATTTAGTCTTGTGTA 1380
Db 64917 ACGAGCGTTGGCTTTATAGAACCTTCTGTGAATTGATCTTGTATTTAGTCTTGTGTA 64976
Oy 1381 GGAAGAAAGAGCAAGTATGCTTGGACCGGATTTAGTGACAAATGATGAGACTTTGGA 1440
Db 64977 GGAAGAAAGAGCAAGTATGCTTGGACCGGATTTAGTGACAAATGATGAGACTTTGGA 65036
Oy 1441 AGAGATCACTCCGCGTTATGCTTGGAGCTACTTGGCTTACCGGTTGAGTGAATTAAGC 1500
Db 65037 AGAGATCACTCCGCGTTATGCTTGGAGCTACTTGGCTTACCGGTTGAGTGAATTAAGC 65096
Oy 1501 TGGCAAAAGACTAAATGCTTTAAGCGGTGTCGGAATATTTTGTGCTGTGAGAGAG 1560
Db 65097 TGGCAAAAGACTAAATGCTTTAAGCGGTGTCGGAATATTTTGTGCTGTGAGAGAG 65156
Oy 1561 TGGAGCAATCAGCTCTTGTGTGGGGGTTTGAACCGGTGAGAAAGTTATGAATGAGCGCTTTT 1620
Db 65157 TGGAGCAATCAGCTCTTGTGTGGGGGTTTGAACCGGTGAGAAAGTTATGAATGAGCGCTTTT 65216
Oy 1621 ACGAATGACAGCTGTGAGAGAGTATACAGTTAGATACCTTTTATTTTATTTTATGACA 1680
Db 65217 ACGAATGACAGCTGTGAGAGAGTATACAGTTAGATACCTTTTATTTTATTTTATGACA 65276
Oy 1681 TGAATTAACCTTAAAGTCTCATTTTAAATGATATGTTGTGTGAGTGTGATCTTTTGTGA 1740
Db 65277 TGAATTAACCTTAAAGTCTCATTTTAAATGATATGTTGTGTGAGTGTGATCTTTTGTGA 65336
Oy 1741 GCTACCCCAAGAAATATCCAGAGAGTCATTTGAAAGTTTACGAAAGTTGACCTTGTCTT 1800
Db 65337 GCTACCCCAAGAAATATCCAGAGAGTCATTTGAAAGTTTACGAAAGTTGACCTTGTCTT 65396
Oy 1801 GTGGCTCAGCTTTTATGTGTAGAAAGCACACTTTTACAGSAGTCTGATPAGCAATTG 1860
Db 65397 GTGGCTCAGCTTTTATGTGTAGAAAGCACACTTTTACAGSAGTCTGATPAGCAATTG 65456
Oy 1861 CAGCAACTTCAGCAGGCTPAGGTAAATGCTATGAGATTCCTGCGATGTTGATGATACA 1920
Db 65457 CAGCAACTTCAGCAGGCTPAGGTAAATGCTATGAGATTCCTGCGATGTTGATGATACA 65516
Oy 1921 CGGAATATTTGGAGATAGACTTCGGCTCTAGAAAGGGACCTGTGACCTGCTTATAGG 1980
Db 65517 CGGAATATTTGGAGATAGACTTCGGCTCTAGAAAGGGACCTGTGACCTGCTTATAGG 65576
Oy 1981 AAAGTTGATGAAATGCGTATGTTGGGCTTGAACAGTGAAGATTCACAATATAGAAAT 2040
Db 65577 AAAGTTGATGAAATGCGTATGTTGGGCTTGAACAGTGAAGATTCACAATATAGAAAT 65636
Oy 2041 CCAAGTATTTGAGAGTTTGTGGAATTCAAATCGTATGACATGATGATCTCCCT 2100
Db 65637 CCAAGTATTTGAGAGTTTGTGGAATTCAAATCGTATGACATGATGATCTCCCT 65696
Oy 2101 GGACATGCAAAATTTGTTGGAACCTGTGTGCAAGGGGTTGTCTTCCATGAGTTGAGAGAC 2160
Db 65697 GGACATGCAAAATTTGTTGGAACCTGTGTGCAAGGGGTTGTCTTCCATGAGTTGAGAGAC 65756
Oy 2161 ACCAAAGATTAATAATTTAACTCGGGGACTACTATGATGATCCATGTTTGGTTAC 2220
Db 65757 ACCAAAGATTAATAATTTAACTCGGGGACTACTATGATGATCCATGTTTGGTTAC 65816
Oy 2221 TTGGAAGAGTGAAGTATGAGGTTCTCTTTTAACTGTCTGCAATGATGCAAGG 2280
Db 65817 TTGGAAGAGTGAAGTATGAGGTTCTCTTTTAACTGTCTGCAATGATGCAAGG 65876
Oy 2281 ATTGAGCCGAGACATGTGAAGCTATGCTATGAGGACATGACAGAAAGTTTCTTCC 2340
Db 65877 ATTGAGCCGAGACATGTGAAGCTATGCTATGAGGACATGACAGAAAGTTTCTTCC 65936

Oy 2341 CGCTATACAGATAGAAACTCGGCTGAACCCAGAGATGTGCAGAGACAGTGTATGTTA 2400
Db 65937 CGCTATACAGATAGAAACTCGGCTGAACCCAGAGATGTGCAGAGACAGTGTATGTTA 65996
Oy 2401 GATCTGTGTGTAACTAATGATGAGCGCTGATGTGAGGCTGTGTCTTTATTTGCAAGACT 2460
Db 65997 GATCTGTGTGTAACTAATGATGAGCGCTGATGTGAGGCTGTGTCTTTATTTGCAAGACT 66056
Oy 2461 GTAGAACCTCTGAAAACCTTGAAACCTAATGATTAATGCAATGAGGCTGAGGCTCAGAG 2520
Db 66057 GTAGAACCTCTGAAAACCTTGAAACCTAATGATTAATGCAATGAGGCTGAGGCTCAGAG 66116
Oy 2521 AGTAGCGTTGATGAACCTAATGATGAAATGTCGTTGCTGATATGTTAAAGAGGACAGT 2580
Db 66117 AGTAGCGTTGATGAACCTAATGATGAAATGTCGTTGCTGATATGTTAAAGAGGACAGT 66176
Oy 2581 GTGAAGATCTGAGCTGTGTGTGTGCAATTTGACATGATTTCACTGTTACCGAGAAAT 2640
Db 66177 GTGAAGATCTGAGCTGTGTGTGTGCAATTTGACATGATTTCACTGTTACCGAGAAAT 66236
Oy 2641 TTTCTTAAAGCAAGCTCATCTTTTCAAGCAAGGATATGTTCTTCTATGAAATCGAT 2700
Db 66237 TTTCTTAAAGCAAGCTCATCTTTTCAAGCAAGGATATGTTCTTCTATGAAATCGAT 66296
Oy 2701 GTCGCTACCATAGGTATGATTAATGATGCAATTTTCAATATCTGATTTGCTCAAAATA 2760
Db 66297 GTCGCTACCATAGGTATGATTAATGATGCAATTTTCAATATCTGATTTGCTCAAAATA 66356
Oy 2761 TGTCTGTTTGTGTAGCTTAAGAACTATGTTCCCATTAATCATGTGCCAAAGTTGATAC 2820
Db 66357 TGTCTGTTTGTGTAGCTTAAGAACTATGTTCCCATTAATCATGTGCCAAAGTTGATAC 66416
Oy 2821 AAGATTAACAAGTTGCTGAGTAATTTCACTAATATGCTGTGAATTTTGTGATCAAA 2880
Db 66417 AAGATTAACAAGTTGCTGAGTAATTTCACTAATATGCTGTGAATTTTGTGATCAAA 66476
Oy 2881 CTGTAGACAGAAATGTAATTTTCACTCTCAACATTTCTGTTTGAATPAGTAGATTAAG 2940
Db 66477 CTGTAGACAGAAATGTAATTTTCACTCTCTCAACATTTCTGTTTGAATPAGTAGATTAAG 66536
Oy 2941 AGATTGCTTATGTTGTGCTTGTGTCAACTTTTCTTCTGATTTTCTTTTCCATTT 3000
Db 66537 AGATTGCTTATGTTGTGCTTGTGTCAACTTTTCTTCTGATTTTCTTTTCCATTT 66596
Oy 3001 AAGGTCAGTACAGAGCTGAGATTCAGAGACATTCACAGAAATGAGTCTAGACCTGACA 3060
Db 66597 AAGGTCAGTACAGAGCTGAGATTCAGAGACATTCACAGAAATGAGTCTAGACCTGACA 66656
Oy 3061 GAATATAGTATCCAGAGTGCAGAAATTAAGTCTGTGCTTTTGGGCTGTGATCACGCAAT 3120
Db 66657 GAATATAGTATCCAGAGTGCAGAAATTAAGTCTGTGCTTTTGGGCTGTGATCACGCAAT 66716
Oy 3121 AGAAATGTTTACAGAGGTGAGGGAATTAATCTACATTAATCAATTTGTGTGAATCTGT 3180
Db 66717 AGAAATGTTTACAGAGGTGAGGGAATTAATCTACATTAATCAATTTGTGTGAATCTGT 66776
Oy 3181 TGGACATGATTAATAGTGTGTGCTGTTGATTTCTGTATTAATAGTTTGGATGGG 3240
Db 66777 TGGACATGATTAATAGTGTGTGCTGTTGATTTCTGTATTAATAGTTTGGATGGG 66836
Oy 3241 GAATGCTGAAGATTTGAGCTGACAGAGCAGCTGAAATCTGCGAGCTTGGGTTGGTTATG 3300
Db 66837 GAATGCTGAAGATTTGAGCTGACAGAGCAGCTGAAATCTGCGAGCTTGGGTTGGTTATG 66896
Oy 3301 ATTAATACACTGTTGAAACTATCTGTTGACAGTGTGACAGTCTGACAGATGAAACCGCTG 3360
Db 66897 ATTAATACACTGTTGAAACTATCTGTTGACAGTGTGACAGTCTGACAGATGAAACCGCTG 66956
Oy 3361 CTGTGTGGAAGCAACTGTGGAAGAGTCTGCTGTATCTGATGATTTGGTTCATCCAGAAA 3420
Db 66957 CTGTGTGGAAGCAACTGTGGAAGAGTCTGCTGTATCTGATGATTTGGTTCATCCAGAAA 67016

QY 3421 ACAATGCTACTGATGTGACAGACCTACACAAAGATGCAAGTTTCTGCTCCAGTCTG 3480
 DB 67017 ACAATGCTACTGATGTGACAGACCTACACAAAGATGCAAGTTTCTGCTCCAGTCTG 67076
 QY 3481 GGTGGAAATCACTGAAGGCTGTTCTTCGATCATATATATCTCATATGTCAGTCTG 3540
 DB 67077 GGTGGAAATCACTGAAGGCTGTTCTTCGATCATATATATCTCATATGTCAGTCTG 67136
 QY 3541 GAGCTTGGAGATTCCTCTTCTTCTGTAATTCCTCTCTTAAGTTAGTTTATTAATGA 3600
 DB 67137 GAGCTTGGAGATTCCTCTTCTTCTGTAATTCCTCTCTTAAGTTAGTTTATTAATGA 67196
 QY 3601 ACACAAAAAATTAACGTTCTTGACACACCCCTTCTTCTGATCTAACTAACTAACTAAG 3660
 DB 67197 ACACAAAAAATTAACGTTCTTGACACACCCCTTCTTCTGATCTAACTAACTAAG 67256
 QY 3661 GCTACAA 3667
 DB 67257 GCTACAA 67263

RESULT 4
 AY091075
 LOCUS
 DEFINITION
 AY091075 2637 bp mRNA linear PLN 18-SEP-2002
 Arabidopsis thaliana unknown protein (At5g42480) mRNA, complete
 cds.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 AY091075
 AY091075.1 GI:20259550
 FLI CDNA.
 Arabidopsis thaliana (thale cress)
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE
 AUTHORS
 1 (bases 1 to 2637)
 Yamada, K., Banh, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M.,
 Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L.,
 Tang, C.C., Tortum, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L.,
 Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,
 Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,
 Liu, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J.,
 Sakurai, T., Satou, M., Seki, M., Shim, P., Southwick, A.,
 Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.
 Arabidopsis Full Length cDNA Clones

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 2 (bases 1 to 2637)
 Yamada, K., Banh, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M.,
 Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L.,
 Tang, C.C., Tortum, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L.,
 Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,
 Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,
 Liu, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J.,
 Sakurai, T., Satou, M., Seki, M., Shim, P., Southwick, A.,
 Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.
 Direct Submission
 Submitted (21-MAR-2002) Plant Gene Expression Center, 800 Buchanan
 Street, Albany, CA 94710, USA
 RIKEN Genomic Sciences Center (GSC) members carried out the
 collection and clustering of RAPL cDNAs (RAPL cDNA: 'RIKEN
 Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J.,
 Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,
 Hayashizaki, Y. and Shinozaki, K.

TITLE
 JOURNAL
 COMMENT
 The Salik, Stanford, PGSC (SSP) Consortium members carried out the
 sequencing and annotation of the RAPL cDNAs: Yamada, K., Banh, J.,
 Chan, M.M., Chang, C.H., Chang, E., Dale, J.M., Deng, J.M.,
 Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C.,
 Tortum, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Chen, H.,
 Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Lam, B., Liu, J.,
 Meyers, M.C., Miranda, M., Nguyen, M., Palm, C.J., Shim, P.,
 Southwick, A., Davis, R.W., Ecker, J.R. and Theologis, A.
 Yamada, K. (SSP/PGSC) and Seki, M. (RIKEN GSC) contributed equally to

FEATURES
 source
 Annotation is based on the January 2002 version of the Arabidopsis
 genome submitted to GenBank.
 Location/Qualifiers
 1..2637
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /chromosome="5"
 /clone="RAF09-76-G11 (R19395)"
 /ecotype="Columbia"
 /note="This clone is in a modified pBluescript vector
 (pLC-1) as a BamHI/XhoI insert."
 1..2637
 /gene="At5g42480"
 1..113
 /gene="At5g42480"
 114..2519
 /gene="At5g42480"
 /codon_start=1
 /evidence=experimental
 /product="unknown protein"
 /protein_id="AAM13895.1"
 /db_xref="GI:20259551"
 /translation="MEALSHVIGIGLSPOLGRLPATTKLRSHNTSTTICSAKMD
 RLSDNFTSDSSSSPATATTTATVATVPSISDPSRHVPIPDVGLAQHFLT
 DGIIRAPARVSKPQGFSDALISRQIIQAACETISNRSRRRTNGLDDEAT
 VITVPMKVPALCVLQEGEETELVGLALKERIPKSFQDVVLMALAFIDVS
 RDNAALDPDFITGEFVEALKLOEGASLAPDLAQIDLEETPRVLELDG
 LPDLDDYAKRLNGISGVRLIMVGGGASLAVGLTRERFNAFLRMTAAQVDL
 FVAPISNPSPSEVEYVALVVAQFICKPHLDADKFOOLQOAKVAMEIIPAM
 LYDRNMIEDIFGELGICALLIGKVCRMWILGDSDSQVRNPAIVFVLENSRD
 DNDLRLGKLETLTLAGVYPPRPDTODKPKGDIYDDPMVLSYLERVVOGSP
 AAAMARIIGAEHYKASMQLVQVFERITDRKSAEPKDVQETVSVDPVGNNGRD
 GEPGFIAEAVRPSNFEPTNDYALRAVSESSVDETVEMSVADMKEASVYKILAGV
 AIGISIFSQYFIFKSSSPQRKDMVSESDVATIGVRADDSALPRMPTAETNI
 VSKQKIKSLAFGPDHRIEMPEVLDGMLKITDRAETQOLGLVDTTLKLSVDS
 VTVASDGRALVEATLLESALSLDVLHENNATVRYTTTYEVWMSKMKITGVS
 LAS"
 2520..2637
 /gene="At5g42480"

gene
 5'UTR
 CDS

ORIGIN
 3'UTR

Query Match 52.7%; Score 1931; DB 15; Length 2637;
 Best Local Similarity 80.4%; Pred. No. 0;
 Matches 2616; Conservative 0; Mismatches 5; Indels 632; Gaps 5;

QY 368 GATTTAACTTACTACTCAATCAAAATCCATTAACCTTAGACGACCAAGAGTCTC 427
 DB 1 GATTTAACTTACTACTCAATCAAAATCCATTAACCTTAGACGACCAAGAGTCTC 60
 QY 428 TTCAATATGTAAACAGAACAAAGTTTGTAGTAGCTTAAGAAGACACTCCATGAG 487
 DB 61 TTCAATATGTAAACAGAACAAAGTTTGTAGTAGCTTAAGAAGACACTCCATGAG 120
 QY 488 CTGAGGACAGTGGGCAATGGCTCCGCCATTCGAATTATGCGAATTACACCGGGA 547
 DB 121 CTGAGGACAGTGGGCAATGGCTCCGCCATTCGAATTATGCGAATTACACCGGGA 180
 QY 548 CGAAGAGCTCGAGCTGACCAACACCTTACAACTATGCTCCGCCAGCAAAATGG 607
 DB 181 CGAAGAGCTCGAGCTGACCAACACCTTACAACTATGCTCCGCCAGCAAAATGG 240
 QY 608 CCGACCGCTTCTCTCGGACTTCAATTTACCTCGATTCCTCTCTCTCTTGGCCA 667
 DB 241 CCGACCGCTTCTCTCGGACTTCAATTTACCTCGATTCCTCTCTCTCTTGGCCA 300
 QY 668 CCGCACACACACCGGCACTGCTCTGCGACCACTATGATGTCGCGAAGGC 727
 DB 301 CCGCACACACACCGGCACTGCTCTGCGACCACTATGATGTCGCGAAGGC 360

QY 728 ACCTCCCATCCCATTTGATTTCTACAGAGTATTAGAGCTCAACACATTTCTTAACG 787
DB 361 ACCCTCCCATCCCATTTGATTTCTACAGAGTATTAGAGCTCAACACATTTCTTAACG 420
QY 788 ATGGAATGAGAAGAGCATTCGAAAGCTAGAGGTTTGAAAACCGCCGAATTCGGTTTCAGCG 847
DB 421 ATGGAATGAGAAGAGCATTCGAAAGCTAGAGGTTTGAAAACCGCCGAATTCGGTTTCAGCG 480
QY 848 ACAGAGCTTTATACAGCCGAGACAGATTTCTTCAAGCTCTTGCAGAACTCTGTCTAATC 907
DB 481 ACAGAGCTTTATACAGCCGAGACAGATTTCTTCAAGCTCTTGCAGAACTCTGTCTAATC 540
QY 908 CTCGGTCTAGAAGAGATTAATGAAGGCTCTTCTGATGATGAAGAGCTACAGTCATCA 967
DB 541 CTCGGTCTAGAAGAGATTAATGAAGGCTCTTCTGATGATGAAGAGAGCTACAGTCATCA 600
QY 968 CTGATGTTCTCTGGATTAAGGTAATTTGGATTTGCGAATAATAAGTTTCTTCGTTTAA 1027
DB 601 CTGATGTTCTCTGGAT----- 617
QY 1028 TTTGATGATTTGGAATAAGAGAACTTTTATCTAGTGAAGGTTCTGCGGCTCTCTGT 1087
DB 618 -----AAGTTTCTGTGCTCTCTGT 638
QY 1088 GTATTTGCAAGAGGCTGTGAGACTGAGATAGTTCTTCCGGTTGTGAGGCTCTGCTTAAG 1147
DB 639 GTATTTGCAAGAGGCTGTGAGACTGAGATAGTTCTTCCGGTTGTGAGGCTCTGCTTAAG 698
QY 1148 GAGAGGTTGCTTAAGTCTGTTTAAGCAAGATGTTGTTTATGTTATGAGGCTGAGCTTCTC 1207
DB 699 GAGAGGTTGCTTAAGTCTGTTTAAGCAAGATGTTGTTTATGTTATGAGGCTGAGCTTCTC 758
QY 1208 GATGCTTCGAGGAGTCTATGAGCAATGATCCACTGATTTTATCTGTTTATGAGTT 1267
DB 759 GATGCTTCGAGGAGTCTATGAGCAATGATCCACTGATTTTATCTGTTTATGAGTT 818
QY 1268 GTTGAAGAGCTTTGAGCTTTTACAGGTAGTTGACTTCTTGTGTAATTTGACGAGCG 1327
DB 819 GTTGAAGAGCTTTGAGCTTTTAC----- 843
QY 1328 TTGGCTTTATAAGAACTTTCTTGATTTGATACTTTGTTATGAGTCTTGTGAGAGGAA 1387
DB 844 -----AAGAGGAA 851
QY 1388 GGAGCAAGTACCTTTGCAACCGGATTTTACGTGCACAAATTGATGAGACTTTGGAAGAGATC 1447
DB 852 GGAGCAAGTACCTTTGCAACCGGATTTTACGTGCACAAATTGATGAGACTTTGGAAGAGATC 911
QY 1448 ACTCCGCTTATGTTCTTGGAGCTACTTGGCTTACCCTTGGTATGATTAAGCTGCGAAA 1507
DB 912 ACTCCGCTTATGTTCTTGGAGCTACTTGGCTTACCCTTGGTATGATTAAGCTGCGAAA 971
QY 1508 AGACTAAATGTTTTAAGCGGTGTGCGAAATTTTGTGTCTGTTGAGAGAGGTGAGACA 1567
DB 972 AGACTAAATGTTTTAAGCGGTGTGCGAAATTTTGTGTCTGTTGAGAGAGGTGAGACA 1031
QY 1568 TCAGCTCTTGTGGGGGTTTGAACCGTGAGAAATTATGAATGAGGCTTTTTCAGAAATG 1627
DB 1032 TCAGCTCTTGTGGGGGTTTGAACCGTGAGAAATTATGAATGAGGCTTTTTCAGAAATG 1091
QY 1628 ACAAGCTGTAAGCAAGTATACAGTTTGAATCCTTTTAAATTTCTTTACAGATGATATA 1687
DB 1092 ACAAGCTGTAAGC----- 1104
QY 1688 ACTTAGGTTTCTCATTTTAATGTAATGTTGTGTGTAAGTGAATCTTTTGTAGTACC 1747
DB 1105 -----AGTTGATCTTTTGTAGTACC 1128
QY 1748 CAAGCAATATTCAGCAGAGTCAATTTGAAGTTTACGAAGTTGCACTTCTCTGTGAGCTC 1807
DB 1129 CAAGCAATATTCAGCAGAGTCAATTTGAAGTTTACGAAGTTGCACTTCTCTGTGAGCTC 1188
QY 1808 AAGCTTTTATTTGTAAGAGCAACCTTTTACAGAGTCTGATTAAGCAATTCAGCAAC 1867

DB 1189 AAGCTTTTATTTGTAAGAACCAACCTTTTACAGAGTCTGATTAAGCAATTCAGCAAC 1248
QY 1868 TTCAGCAGGCTAAGAGTATGAGCTATGAGAGTTCCTGCGATGTTGTATGATACACGAAATA 1927
DB 1249 TTCAGCAGGCTAAGAGTATGAGCTATGAGAGTTCCTGCGATGTTGTATGATACACGAAATA 1308
QY 1928 ATTGGAGATAGACTTCGCTCTAGAAAAGGGGACTCTGTGCACCTGCTTATAGGCAAGCTTG 1987
DB 1309 ATTGGAGATAGACTTCGCTCTAGAAAAGGGGACTCTGTGCACCTGCTTATAGGCAAGCTTG 1368
QY 1988 ATGAATCCGCTATGATGTTGGGCTTACAGTGAAGATTCACAAATAGGAATCCAGCTA 2047
DB 1369 ATGAATCCGCTATGATGTTGGGCTTACAGTGAAGATTCACAAATAGGAATCCAGCTA 1428
QY 2048 TTGTGAGATTTGTTTGAAGAAATTCAAATGCTGATGACAAATGATATCTCCCTGAGCTAT 2107
DB 1429 TTGTGAGATTTGTTTGAAGAAATTCAAATGCTGATGACAAATGATATCTCCCTGAGCTAT 1488
QY 2108 GCAAAATTTGGAACCTGTGTGCAAGGGTGTCTTCTCTAGGTTGAGAGACCAAAAG 2167
DB 1489 GCAAAATTTGGAACCTGTGTGCAAGGGTGTCTTCTCTAGGTTGAGAGACCAAAAG 1548
QY 2168 ATAAAAATTTAAACTCGGGGACTACTATGATATCTATAGTTTTGAATTAATCTGAAA 2227
DB 1549 ATAAAAATTTAAACTCGGGGACTACTATGATATCTATAGTTTTGAATTAATCTGAAA 1608
QY 2228 GAGTGAAGTATGTTCAAGGTTCTCTCTTACGTGTCGCAACTATAGGCAAGATTTGAG 2287
DB 1609 GAGTGAAGTATGTTCAAGGTTCTCTCTTACGTGTCGCAACTATAGGCAAGATTTGAG 1668
QY 2288 CCGAGCATGTGAAGACTAGTGTATGAGCAAGCTGCAAAAAGTTTTCCTTCCGCTATA 2347
DB 1669 CCGAGCATGTGAAGACTAGTGTATGAGCAAGCTGCAAAAAGTTTTCCTTCCGCTATA 1728
QY 2348 CAGATTAAGAACTCGGCTGAAACCCAAAGATGTTGCAAGAGAGTATAGTATGATCTCTG 2407
DB 1729 CAGATTAAGAACTCGGCTGAAACCCAAAGATGTTGCAAGAGAGTATAGTATGATCTCTG 1788
QY 2408 TTGCTAACAAATGTAAGCCGTGATGTTGAGCTGAGTCTTATTTGCAAGACTGTAAAGAC 2467
DB 1789 TTGCTAACAAATGTAAGCCGTGATGTTGAGCTGAGTCTTATTTGCAAGACTGTAAAGAC 1848
QY 2468 CTTCTGAAAACCTTGAACCTATGATTAATGATTAATGCAATTTGAGCTGGGCTCTCAGAGATAGCG 2527
DB 1849 CTTCTGAAAACCTTGAACCTATGATTAATGATTAATGCAATTTGAGCTGGGCTCTCAGAGATAGCG 1908
QY 2528 TTGATGAACCTACTGTTGAAGATGTCCTGTGCTGATGTTTAAAGAGGCAAGGTGAAGA 2587
DB 1909 TTGATGAACCTACTGTTGAAGATGTCCTGTGCTGATGTTTAAAGAGGCAAGGTGAAGA 1968
QY 2588 TCTAGGCTGCTGTGTGGAATTTGCACTGATTTCACTGTTCAAGCAGAAAGTATTTTCTTA 2647
DB 1969 TCTAGGCTGCTGTGTGGAATTTGCACTGATTTCACTGTTCAAGCAGAAAGTATTTTCTTA 2028
QY 2648 AAGCAGCTCATCTTTTCAACGCAAGATAGTGTCTTCTATGAAATCTGATGTGCTTA 2707
DB 2029 AAGCAGCTCATCTTTTCAACGCAAGATAGTGTCTTCTATGAAATCTGATGTGCTTA 2088
QY 2708 CCAATAGTATGATTAATGATGCAATTTTCAATATCTGCAATTCCTCAAAATATGCTTGT 2767
DB 2089 CCA----- 2091
QY 2768 TTTGTGAGTAAAGAACATAGTTCCTAATAACATGTCCAAAAGTTGTACCAAGATTA 2827
DB 2092 ----- 2091
QY 2828 ACAAGTGTCTGAGTAAATTTCACTAATTAATGCTGCTGAAATTTTGTGATCAAACTGTA 2887
DB 2092 ----- 2091
QY 2888 CAGAAATGTAATTTTCACTCTCAACATTTCTGTTTGAATTAAGGTAGATTAAGATTTGC 2947

Db	2092	-----	2092	-----
Qy	2948	CTTAGTGTGGCTTTGTGCCAATTCTTTCTTCTTGATTTTTTTCTTTGCATTTAGAGTGCA	3007	-----
Db	2092	-----	2092	-----
Qy	3008	GTCAGAGCTGACGATTGCAAGGACCTTCCGAAATGATGCTAGGACCTGCAGAGAAATTA	3067	-----
Db	2100	GTCCAGGCTGACGATTGCAAGGACCTTCCGAAATGATGCTAGGACCTGCAGAGAAATTA	2155	-----
Qy	3068	GTATCCAAAGTGGCAGAAAGATTAACTCTCTGCTTTTGGGCGCTGATCACCAGATAGAATG	3127	-----
Db	2160	GTATCCAAAGTGGCAGAAAGATTAACTCTCTGCTTTTGGGCGCTGATCACCAGATAGAATG	2219	-----
Qy	3128	TTAACCAAGGTGAGGGAATTAATTCACAATTCAAATGATGTAAGAACTGTGGACAT	3187	-----
Db	2220	TTAACCAAGGTGAGGGAATTAATTCACAATTCAAATGATGTAAGAACTGTGGACAT	2226	-----
Qy	3188	GATTATAGTCTGGTGCCCTTGTGTTGATTTCTGTTATTAAGGTTTGGATGGCCGAATGCT	3247	-----
Db	2227	-----	2248	-----
Qy	3248	GAGATTTTGAAGCTGCAGAGAGCTGAAACTGCCAGCTTGGGTTGTTATTAATATAC	3307	-----
Db	2249	GAGATTTTGAAGCTGCAGAGAGCTGAAACTGCCAGCTTGGGTTGTTATTAATATAC	2308	-----
Qy	3308	ACTGTTGAAACTATCTGTTTGAACAGTGTCACAGTCTCAGCAGATGGAACCCGTGCTTGCT	3367	-----
Db	2309	ACTGTTGAAACTATCTGTTTGAACAGTGTCACAGTCTCAGCAGATGGAACCCGTGCTTGCT	2368	-----
Qy	3368	GGAAGCAACTCTGGAGAGTCTGCTGTTCTATCTGATTTGTTGTTCAACAGAAACAATGC	3427	-----
Db	2369	GGAAGCAACTCTGGAGAGTCTGCTGTTCTATCTGATTTGTTGTTCAACAGAAACAATGC	2428	-----
Qy	3428	TACGATGTCGAAAGCTTACACAAACAAGTATAGAAATTTTCGTGCCAAGTCAGAGGTGTA	3487	-----
Db	2429	TACGATGTCGAAAGCTTACACAAACAAGTATAGAAATTTTCGTGCCAAGTCAGAGGTGTA	2488	-----
Qy	3488	AATCACTGAAGGCTCTGTTCTTGATCATCAATTAATCAATATGATGATGATGATGATGATG	3547	-----
Db	2489	AATCACTGAAGGCTCTGTTCTTGATCATCAATTAATCAATATGATGATGATGATGATGATG	2548	-----
Qy	3548	CGAATTCCTTTGTTCTGTAATTCCTCTCTCTAAATTAATGATGATTAATTAATGAACACAA	3607	-----
Db	2549	CGAATTCCTTTGTTGTTGTAATTCCTCTCTCTAAATTAATGATGATTAATTAATGAACACAA	2608	-----
Qy	3608	AAAATTAACCTTC 3620	-----	-----
Db	2609	AAAATTAACCTTC 2621	-----	-----
RESULT 5	AY221469	2438 bp	linear	PLN 05-AUG-2003
LOCUS	AY221469	2438 bp	linear	PLN 05-AUG-2003
DEFINITION	Arabidopsis thaliana division protein (AR6) mRNA, complete cds;			
ACCESSION	AY221469			
VERSION	AY221469.1	GI:33436353		
KEYWORDS				
SOURCE	Arabidopsis thaliana (thale cress)			
ORGANISM	Arabidopsis thaliana			
REFERENCE				
AUTHORS	Viltha,S., Froehlich,J.E., Kosharova,O., Pyke,K.A., Van Ert,H. and Oosteryoung,K.W.			
TITLE	AR6 is a J-domain Plastid Division Protein and an Evolutionary			
JOURNAL	Descendant of the Cyanobacterial Cell Division protein Ftn2			
PUBLISHED	Plant Cell 15 (8), 1918-1933 (2003)			
FIGURE	12897262			
AUTHORS	2 (bases 1 to 2438)			
	Viltha,S., Kosharova,O., van Ert,H., Froehlich,J.E. and			

TITLE		Osteryoung, K. M.	
JOURNAL		Direct Submission	
FEATURES		Submitted (17-JAN-2003) Department of Plant Biology, Michigan State University, 166 Plant Biology Building, East Lansing, MI 48824, USA	
source		location/Qualifiers	
1. 2438		/organism="Arabidopsis thaliana"	
		/mol_type="mRNA"	
		/db_xref="taxon:3702"	
		/ecotype="wassilewskija"	
gene		1. 2438	
		/gene="ARC6"	
		/note="synonym: At5g42480"	
CDS		2. 2407	
		/gene="ARC6"	
		/codon_start=1	
		/product="division protein"	
		/protein_id="AA018646.1"	
		/db_xref="GI:33436354"	
		/translation="MEALSHVGIQLSPQLCLPPLATTKLRHSNTSTTICSAKWDRLSDPNFTSDSSSSPATATTTATVAVSLPPSIDRPNRPIDIPQVLAGQTHPLTGDLRAPAEARVSPQPFSDADILSRQLDACEPLSRRRERENLEBDEGATVINDPMDKVPKALCVLQEGEGMEIVLRVEALIKRLPSFDQVVLWALAFLDVS RDMALDPPDPTGYEVEBEALGLDQEGASSLAPDLRAQIDISLBITRYRYLLGLPGLGDDYAAKRLNGLSGVRNITLMSVGGGASALVGLTRKKFMPEARLRNTABQVQLFVATNPISPFSEFEVEVALVAQAIFIGKPHLLDADQFQOLQAKWAEIPIAMLYPTRNWMEIDFGLERGLCALIGKVDCEKMLGLDSEDSQYNPAIVEVNVGSSND DNDNDPLCLGLFTMLAGVVEPRPRDCKKFKLGDYDDPMYLSYERVENVGSPL AAAATMARIGAEHYKASAMQALQKVPSPRYTIDNSAPKQVQTFPEVDVGNVNGSD GEGGVFLAEVVRPSENFTNDYAIRAGVSSSSVDTEVENSVAEMLEKASVKILAAQV AIGILSLFSQKFLPKSSSSFORQDOWSMSSDVATIGSVRADSEALPRDARTAEIVSFKWKIKSLAFGRHRIEMLPVLGDRMLKIWTDRALETQAQLGVDTYTLKLSVDS VTVASQGTALVATLEESACLSDLVHPENNATDVRYTYRVEVFMKSGMKITGSGV LAS"	
misc_difference		474	
		/gene="ARC6"	
		/note="results in gly to glu substitution"	
misc_difference		1831	
		/gene="ARC6"	
		/note="results in glu to asp substitution"	
		/replace="t"	
ORIGIN			
Query Match		47.8%; Score 1752.8; DB 15; Length 2438;	
Best Local Similarity		79.3%; Pred. No. 0;	
Matches 2436; Conservative		0; Mismatches 2; Indels 632; Gaps 5;	
QY	480 CATGGAAGCTCTGAGTACGCTGGCACTTGCTCTCTCCCAATTCGATTCGCGATTAC	539	
DB	1 CATGGAAGCTCTGAGTACGCTGGCACTTGCTCTCTCCCAATTCGATTCGCGATTAC	60	
QY	540 ACCGGCGACGACAAAGCTCCGAGGTAGCCCAACAACTCTTACAACTATCTGCTCCGCGAG	599	
DB	61 ACCGGCGACGACAAAGCTCCGAGGTAGCCCAACAACTCTTACAACTATCTGCTCCGCGAG	120	
QY	600 CAATGGGCGGACCGCTTCTCTCCGACTTCAATTCACCTCCGATTCCCTCCCTCTCTC	659	
DB	121 CAATGGGCGGACCGCTTCTCTCCGACTTCAATTCACCTCCGATTCTCTCTCTCTCTC	180	
QY	660 CTTCGCAACGCGACACCAACCGGCACCTCTGCTCTCTGCGACCACTTATTTAGATGCTCC	719	
DB	181 CTTCGCAACGCGACACCAACCGGCACCTCTGCTCTCTGCGACCACTTATTTAGATGCTCC	240	
QY	720 CGAAGCGCACGTCCTCCATCCCATTTGATTTCTACCAAGTATTTAGAGCTCAAAACATTT	779	
DB	241 CGAAGCGCACGTCCTCCATCCCATTTGATTTCTACCAAGTATTTAGAGCTCAAAACATTT	300	
QY	780 CTTAAACCGAGGAAATGAGAGACATTTGAAAGCTAGAGTTTGAACCCGCGGACATTCGG	839	
DB	301 CTTAAACCGAGGAAATGAGAGACATTTGAAAGCTAGAGTTTGAACCCGCGGACATTCGG	360	
QY	840 TTTCAGGACGACGCTTTATTCAGCCGCGAGACAGATTTCTTCAAGCTGCTTGCAGAACTCT	899	

|||||
Db 361 TTTGAGCAGCAGCGCTTAAATCAGCCGAGACAGATTCCTCAAGCTGCTCGAAACTCT 420
Oy 900 GTCTAAATCTGGGTAGAAAGAGTAGACAATGAAGTCTTCTGTAGTGTGAAGAAGCTAC 959
Db 421 GTCTAAATCTGGGTAGAAAGAGTAGACAATGAAGTCTTCTGTAGTGTGAAGAAGCTAC 480
Oy 960 AGTCATCACTAGTCTTCTGGGATTAAGTAAATTTTCAATTCGGAATATTAAGTTTCTT 1019
Db 481 AGTATCACTAGTCTTCTGGGAT----- 505
Oy 1020 CGTTTAAATTCATGAATTGGATTAAGGAAGAACTTTATCTAGTGAAGTTCTCGGG 1079
Db 506 -----AAGTTCTCGGG 518
Oy 1080 CTCTCTGTGTATTTGCAAGAGGTGTGAAGCTGAGATAGTTCTTCGGGTGTGTGAGCTC 1139
Db 519 CTCTCTGTGTATTTGCAAGAGGTGTGTGAAGCTGAGATAGTTCTTCGGGTGTGTGAGCTC 578
Oy 1140 TGTCTTAAGAGAGGTTGCTTAAGTCTGTTTAAAGCAAGATGTGGTTTATGTTATGGCGCTTG 1139
Db 579 TGTCTTAAGAGAGGTTGCTTAAGTCTGTTTAAAGCAAGATGTGGTTTATGTTATGGCGCTTG 638
Oy 1200 CGTTTCTGATGTCTCGAGGAGTGTATGCAATTCACCTGATTTTATCTAGTGT 1259
Db 639 CGTTTCTGATGTCTCGAGGAGTGTCTATGCAATTCACCTGATTTTATCTAGTGT 698
Oy 1260 ATGAGTTTGTGAGGAGCTTTGAAGCTTTTACAGTAGTTTGACTTGCTTGGTAATTT 1319
Db 699 ATGAGTTTGTGAGGAGCTTTGAAGCTTTTAC----- 731
Oy 1320 GACGAGCGTTGGCTTTATTAAGAACTTTCTGTGATTTGATTAATTTGAGTCTTGTGT 1379
Db 732 ----- 731
Oy 1380 AGAGAGAAAGAGCAAGTACCTTTCACCGGATTTACGTGCACAAATTTGATGAGCTTTGG 1439
Db 732 AGAGAGAAAGAGCAAGTACCTTTCACCGGATTTACGTGCACAAATTTGATGAGCTTTGG 791
Oy 1440 AAGAGATCACTCCGGGTATATGCTTGAAGCTACCTTGGCTTACCGCTTGTGTATGATTAAG 1499
Db 792 AAGAGATCACTCCGGGTATATGCTTGAAGCTACCTTGGCTTACCGCTTGTGTATGATTAAG 851
Oy 1500 CTGCGAAAGAATAAATGTTTAAAGCGGTGTGCGGAATATTTTGTGTCTGTTGAGAG 1559
Db 852 CTGCGAAAGAATAAATGTTTAAAGCGGTGTGCGGAATATTTTGTGTCTGTTGAGAG 911
Oy 1560 GTGAGCATCAAGCTTGTGTGGGGTTTGAACCGTGAAGTTTATGATGAGCGCTTTT 1619
Db 912 GTGAGCATCAAGCTTGTGTGGGGTTTGAACCGTGAAGTTTATGATGAGCGCGCTTTT 971
Oy 1620 TACGAATGACAGCTGTGAGCAGTATACAGTTTGAATACCTTTTAAATTTCTTAGC 1679
Db 972 TACGAATGACAGCTGTGAGC----- 992
Oy 1680 ATGATATACTTAGTTCCTCAATTTTATGTATGTGTGTAGTGAATCTTTTGT 1739
Db 993 -----AGTTGATCTTTTGT 1008
Oy 1740 AGCTTACCCCAAGCAATATTCAGCAGATCATTTGAAGTTTACGAAGTTGCACTTGTCT 1799
Db 1009 AGCTTACCCCAAGCAATATTCAGCAGATCATTTGAAGTTTACGAAGTTGCACTTGTCT 1068
Oy 1800 TGTGCTCAAGCTTTTATTTGTGAAGAGCACACTTTTACAGAGTGTGAATTAAGCAAT 1859
Db 1069 TGTGCTCAAGCTTTTATTTGTGAAGAGCACACTTTTACAGAGTGTGAATTAAGCAAT 1128
Oy 1860 CCAGCACTTGAAGAGCTTAAGTAATGCTATGAGATTCCTCGAGATTTGTATGATAC 1919
Db 1129 CCAGCACTTGAAGAGCTTAAGTAATGCTATGAGATTCCTCGAGATTTGTATGATAC 1188
Oy 1920 ACGGAATTAATGGGATTAAGATTCCTGCTTAGAAAGGGGACTCTGTGCACTGCTTAATGG 1979
|||||

Db 1189 ACGGAATTAATGGGATTAAGATTCCTGCTTAGAAAGGGGACTCTGTGCACTGCTTAATAG 1248
Oy 1980 CAAGTTGATGAATGCCGTATGTGTTGGCTTAGACAGTGAAGGATTCACAAATATAGAA 2039
Db 1249 CAAGTTGATGAATGCCGTATGTGTTGGCTTAGACAGTGAAGGATTCACAAATATAGAA 1308
Oy 2040 TCCAGCTAATTTGGAGTTTGTTTTGGAGAAATTCAAATCGTATGACAAATGATCTCCC 2099
Db 1309 TCCAGCTAATTTGGAGTTTGTTTTGGAGAAATTCAAATCGTATGACAAATGATCTCCC 1368
Oy 2100 TGGACTATGCAAAATTTTGGAGAAACCTGTGTGGAGGGGTGTCTTCTAGTTTCAAGA 2159
Db 1369 TGGACTATGCAAAATTTTGGAGAAACCTGTGTGGAGGGGTGTCTTCTAGTTTCAAGA 1428
Oy 2160 CACCAAGATTAATAAATTTTAACTCGGGGACTATATGATGATCTATGTTTGAATTA 2219
Db 1429 CACCAAGATTAATAAATTTTAACTCGGGGACTATATGATGATCTATGTTTGAATTA 1488
Oy 2220 CTTGGAAGAGTGAAGTGAAGTTCAGGGTTCTCTTATGCTGTGCTGCAACTATGCAAG 2279
Db 1489 CTTGGAAGAGTGAAGTGAAGTTCAGGGTTCTCTTATGCTGTGCTGCAACTATGCAAG 1548
Oy 2280 GATTGAGCCGAGCATGTGAAGTGAAGTGTATGCTATGCAAGCACTGCAAGATTTTCTTC 2339
Db 1549 GATTGAGCCGAGCATGTGAAGTGAAGTGTATGCTATGCAAGCACTGCAAGATTTTCTTC 1608
Oy 2340 CCGCTATACAGATTAAGAACTCGGCTGAACCAAGATGTGCAAGAGCAGTGTTTAGTGT 2399
Db 1609 CCGCTATACAGATTAAGAACTCGGCTGAACCAAGATGTGCAAGAGCAGTGTTTAGTGT 1668
Oy 2400 AGATCTGTGTGTGAACAATGTAAGTGAAGTGTGAGCTGTGTCTTTATTTGCAAGC 2459
Db 1669 AGATCTGTGTGTGAACAATGTAAGTGAAGTGTGAGCTGTGTCTTTATTTGCAAGC 1728
Oy 2460 TGTGAAGCCCTCGAAACCTTTGAACCTAATGATTAATGCAATTCGAGCTGGGCTGAGA 2519
Db 1729 TGTGAAGCCCTCGAAACCTTTGAACCTAATGATTAATGCAATTCGAGCTGGGCTGAGA 1788
Oy 2520 GAGTAGCGTGAATGAACCTACTGTTGAATGTCCGTGTGATATGTTTAAAGAGGCAAG 2579
Db 1789 GAGTAGCGTGAATGAACCTACTGTTGAATGTCCGTGTGATATGTTTAAAGAGGCAAG 1848
Oy 2580 TGTGAAGTCTTACCTGCTGTGTGTGCAATTTGACATGATTTCACTGTTCAAGCCAGAACTA 2639
Db 1849 TGTGAAGTCTTACCTGCTGTGTGTGCAATTTGACATGATTTCACTGTTCAAGCCAGAACTA 1908
Oy 2640 TTTTCTTAAAGAGCTCACTTTTCAACCGAAGATATGTTTCTTATGAAATCTGA 2699
Db 1909 TTTTCTTAAAGAGCTCACTTTTCAACCGAAGATATGTTTCTTATGAAATCTGA 1968
Oy 2700 TGTGCTAACCATAGTATGATTAATATGATGCAATTTTCAATATCTGATTTGCAAAAT 2759
Db 1969 TGTGCTAACCA----- 1979
Oy 2760 ATGCTTGTTTGTGAGCTTAAGAACATATGTTCCACTTAATATCATGTCCAAAGTTGTAC 2819
Db 1980 ----- 1979
Oy 2820 CAAGATTAACAATTTGTGAGTAATAATTTCACTAATATATGCTTGAATTTTGTATCA 2879
Db 1980 ----- 1979
Oy 2880 ACTGTAGACAGAAATGTAAATTTTCACTGTCAACATTTCTGTTTAAAGATTAAGATTA 2939
Db 1980 ----- 1979
Oy 2940 GAGATTCCTTAATGTGTGCTTTGTCCAACTTTCTTCTTGAATTTTCTTTTCCATTT 2999
Db 1980 ----- 1979
Oy 3000 TAGGTCAGTCAGAGCTGACGATTCAGAAAGCACTTCCAGAAATGAGTGTAGAGCTGAG 3059
Db 1980 TAGGTCAGTCAGAGCTGACGATTCAGAAAGCACTTCCAGAAATGAGTGTAGAGCTGAG 2039
|||||

QY 3060 AGAATATAGTATCCAGTGGCAGAAAGATTAACTCTTGGCTTTGGCCCTGATCAACCGCA 3119
DB 2040 AGAATATAGTATCCAGTGGCAGAAAGATTAACTCTTGGCTTTGGCCCTGATCAACCGCA 2099
QY 3120 TAGAATGTACCAAGGAGGAGGAATTAATCAATTCATCAATTTGTGAAAACTG 3179
DB 2100 TAGAATGTACCAAGGAGGAGGAATTAATCAATTCATCAATTTGTGAAAACTG 2114
QY 3180 TTGACATGATATATAGTGTGCTGCTGCTTTGTTGATTTATTAAGTTTGGATGGG 3239
DB 2115 -----AGGTTTGGATGGG 2128
QY 3240 CGAATGCTGAAGATTGGAATGACAGAGCAGTGAACCTGCGCAGCTTGGGCTTGTAT 3299
DB 2129 CGAATGCTGAAGATTGGAATGACAGAGCAGTGAACCTGCGCAGCTTGGGCTTGTAT 2188
QY 3300 GATTATACACTGTGTAACCTATCTGTTGACAGTGTGACAGTGTGACAGATGGAACCCGT 3359
DB 2189 GATTATACACTGTGTAACCTATCTGTTGACAGTGTGACAGTGTGACAGATGGAACCCGT 2248
QY 3360 GCTCTGTGGAAGCAACTGTGAGAGAGTCTGCTTGTCTATCTGATTTGGTTCACAGA 3419
DB 2249 GCTCTGTGGAAGCAACTGTGAGAGAGTCTGCTTGTCTATCTGATTTGGTTCACAGA 2308
QY 3420 AACAAATGCTACTGATGTGACAACTTACCAACAAGATGTTTGTGTCACAGTCA 3479
DB 2309 AACAAATGCTACTGATGTGACAACTTACCAACAAGATGTTTGTGTCACAGTCA 2368
QY 3480 GGGTGGAAATATCATCTGAGAGGCTCTGTTCTTGATCATTAATTAATCAATTTGAGCATGTC 3539
DB 2369 GGGTGGAAATATCATCTGAGAGGCTCTGTTCTTGATCATTAATTAATCAATTTGAGCATGTC 2428
QY 3540 TGAGCTTGGC 3549
DB 2429 TGAGCTTGGC 2438

RESULT 6
AY150490
LOCUS
DEFINITION Arabidopsis thaliana unknown protein (At5g42480) mRNA, complete cds.
ACCESSION AY150490
VERSION AY150490.1
KEYWORDS GI:23297711
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
REFERENCE 1 (bases 1 to 2436)
AUTHORS Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Huan,V.W., Lee,J.M., Quach,H.L., Tang,C.C., Toriumi,M., Wallender,E.K., Wong,C., Yu,G., Yuen,S., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J., Natsuka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shimizu,P., Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
TITLE Arabidopsis Open Reading Frame (ORF) Clones
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2436)
AUTHORS Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Huan,V.W., Lee,J.M., Quach,H.L., Tang,C.C., Toriumi,M., Wallender,E.K., Wong,C., Yu,G., Yuen,S., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J., Natsuka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shimizu,P., Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
TITLE Submitted (13-SEP-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA
COMMENT The RIKEN Genomic Sciences Center (GSC) members carried out the

collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-length cDNA') : Seki,M., Natsuka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai, J., Hayashizaki,Y. and Shinozaki,K.
The Salk, Stanford, PGC (SSP) Consortium members constructed and sequenced the PUNI (ORF) clones using the RAFL cDNAs: Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Huan,V.W., Lee,J.M., Quach,H.L., Tang,C.C., Toriumi,M., Wallender,E.K., Wong,C., Yu,G., Yuen,S., Carninci,P., Chen,H., Cheuk,R., Jones,T., Kim,C.J., Nguyen,M., Palm,C.J., Shimizu,P., Southwick,A., Tripp,M.G., Wu,T., Davis,R.W., Ecker,J.R. and Theologis,A.
Yamada,K. (SSP/PGC) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP/PGC) contributed equally to this work as PIs.
Annotation is based on the January 2002 version of the Arabidopsis genome submitted to Genbank.
FEATURES
source
1..2436
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="5"
/clone="U19395"
/ecotype="Columbia"
/note="This clone is in PUNI 51."
1..2436
/gene="At5g42480"
1..2406
/gene="At5g42480"
/codon_start=1
/evidence="experimental"
/product="unknown protein"
/protein_id="AA12907.1"
/db_xref="GI:23297712"
/translation="MEALSHVIGLSPEQLRLPPTATTKLRSHNTSFTTCSASKMADRLSDPNTSDSSSSSPATATATATVLPSPSIDRPEHNVIPIDFYVLGQTFELTGIRAFARVSKPQGFSDALISROLOAACEITLNSRPSREVEGLDDEATVTTPMDKVRGALCVLOEGGETIVAVRGALKERLPSKPDQVTVMLALPIDSVDMAALDPPDITITGEPYEBALKILQESGASLAPDLBAQIDETBEITTPYVLELGLPLDGDYAAKRLNGSLVGRNLTAVSGGASALVGLTRKFMNEAFIRMTAEQVDLFPATPSNPASFEYEVALLVAQAFIGKPHILODQFOQLOQAKVAMEIPAMLYOTRNMEIDPLERGICALLIGVDEKCMGLSDSDSOYRPAIVFVLSNRDNDLPGICXKLETLGAVPPRPRTDKKFGIDYDDPMVSYLERVYVQSSPLAAAAAMARIGAEHYKASMOLOKYPFRYDRSABEKDYQETFPSPDPPANNGRDGEPEVFLAEAVRPSBENFTNDYAIRAGVSSVDETVKSVADMLKASVKTILAGV AIGISLFSQRYFLPKSSSFQKDMVSMESDVATIGSRADDSBALPMARVAENI VSKWQIKSLAFGPDHRIEMPEVLIDGMLKMTDRRAETQGLGLVYDTTLKLSVDS VTSVADGTRALVEATLEBSACLSLVHPENNATDVRTYTRVEVWFSKWKITGSAV LAS"
3'UTR
2407..2436
/gene="At5g42480"
misc_difference
2433..2434
/gene="At5g42480"
/note="compared to genomic sequence"
/replace="tlg"
ORIGIN
Query Match 47.6%; Score 1744.6; DB 15; Length 2436;
Best Local Similarity 79.2%; Pred. No. 0;
Matches 2429; Conservative 0; Mismatches 4; Indels 632; Gaps 5;
QY 481 ATGAAGCTGAGAGCAGTGGCTGCTCCCATTCATTAATGCGGATTACCA 540
DB 1 ATGAAGCTGAGAGCAGTGGCTGCTCCCATTCATTAATGCGGATTACCA 60
QY 541 CCGGAGCAAGAAAGCTCCGAGTACCAACAACCTTACAACTATCTGCTCCGACG 600
DB 61 CCGGAGCAAGAAAGCTCCGAGTACCAACAACCTTACAACTATCTGCTCCGACG 120
QY 601 AATGGGCGACCGCTTCTCTCCGACTTAATTCACCTCGATTCTCTCTCTCC 660

Db 121 AAATGGCCGACCGCTTCTCTCCGATTCATTTACCTCGATTCCTCCCTCCCTCC 180
Qy 661 TTGCGCACCGCACACCGACCGCACTGCTGCTCTGCGCACCACTGATTTGATGCTCC 720
Db 181 TTGCGCACCGCACACCGACCGCACTGCTGCTCTGCGCACCACTGATTTGATGCTCC 240
Qy 721 GAACGCGACGTCGCCCATCCCAATTGATTTCTACGAGTATTAGAGCTCAACACATTTG 780
Db 241 GAACGCGACGTCGCCCATCCCAATTGATTTCTACGAGTATTAGAGCTCAACACATTTG 300
Qy 781 TTAAACCGATGGAATTGAAGAGCATTTGAGAGTATGAGTTTGAAACCGCCCGCAATTCGGT 840
Db 301 TTAAACCGATGGAATTGAAGAGCATTTGAGAGTATGAGTTTGAAACCGCCCGCAATTCGGT 360
Qy 841 TTACGCGACGACGCTTTATGACCGGAGACAGATTTCTTCAAGCTGCTTGCGAACTCTG 900
Db 361 TTACGCGACGACGCTTTATGACCGGAGACAGATTTCTTCAAGCTGCTTGCGAACTCTG 420
Qy 901 TCTAATCCTCGGCTTGAAGAGAGTACAATGAGGTCCTTCTGATGATGAAGAACCTACA 960
Db 421 TCTAATCCTCGGCTTGAAGAGAGTACAATGAGGTCCTTCTGATGATGAAGAACCTACA 480
Qy 961 GTCATCACTGATGTCCTTGGGATTAAGGTAATTCGATTCGGAATAATAAGTTTCTTC 1020
Db 481 GTCATCACTGATGTCCTTGGGATTAAGGTAATTCGATTCGGAATAATAAGTTTCTTC 504
Qy 1021 GTTTTAATTTGATGAATTGATAAGAAAGAACTTTATCTAGTGAAGGTTCCGCGGC 1080
Db 505 -----AAGGTTCTGAGTC 518
Qy 1081 TCTCTGTGTAATTCGAAGAGGTGTGAGACTGAGATAGTCTTCCGGTGTGAGGCTCT 1140
Db 519 TCTCTGTGTAATTCGAAGAGGTGTGAGACTGAGATAGTCTTCCGGTGTGAGGCTCT 578
Qy 1141 GCTTAAGAGAGGTGCTTAAGTCTTTAAGCAAAATGTTGGTTTAAAGTAAAGGCTTGC 1200
Db 579 GCTTAAGAGAGGTGCTTAAGTCTTTAAGCAAAATGTTGGTTTAAAGTAAAGGCTTGC 638
Qy 1201 GTTTCGATGTCGAGGAGATGCTATGAGCATTTGACACTGATTTTAAATTAATGATTA 1260
Db 639 GTTTCGATGTCGAGGAGATGCTATGAGCATTTGACACTGATTTTAAATTAATGATTA 698
Qy 1261 TGAAGTTTGTGAAGAGCTTTGAAGCTTTTACAGGTAGTTGACTTGGTAAATTTG 1320
Db 699 TGAAGTTTGTGAAGAGCTTTGAAGCTTTTAC----- 730
Qy 1321 ACGAGCGTTGGCTTTAATAAGAACTTTCTTGATTTGATTAATTGAGCTTGTGTA 1380
Db 731 -----A 731
Qy 1381 GAGAGAAAGAGCAATAGCTTTCACCGGATTTACGTGCACAAAATTGATGAGACTTTG 1440
Db 732 GAGAGAAAGAGCAATAGCTTTCACCGGATTTACGTGCACAAAATTGATGAGACTTTG 791
Qy 1441 AGAGATCACTCCGCTTATGTCCTTGAGACTACTTGCTTACCGCTGCTGATGATTAAGC 1500
Db 792 AGAGATCACTCCGCTTATGTCCTTGAGACTACTTGCTTACCGCTGCTGATGATTAAGC 851
Qy 1501 TGGCAAAAGACTAATGCTTTAAGCGGTGTGCGGAATTTTGTGTGCTGTTGAGAGAG 1560
Db 852 TGGCAAAAGACTAATGCTTTAAGCGGTGTGCGGAATTTTGTGTGCTGTTGAGAGAG 911
Qy 1561 TGGAGCATCAGCTGTTGTGGGGGTTTGACCGGTGAGAAAGTTTAAAGAAAGGCGTTT 1620
Db 912 TGGAGCATCAGCTGTTGTGGGGGTTTGACCGGTGAGAAAGTTTAAAGAAAGGCGTTT 971
Qy 1621 ACGATGACAGCTGCTGACAGGTATACAGTTAGATACTTTTATTTATTTCTTTAGCA 1680
Db 972 ACGATGACAGCTGCTGAC----- 991
Qy 1681 TGATATTAACCTTAGGTTTCTCATTTTAAATGATATGTTGTGTAGGTTGATCTTTTGT 1740
Db 1740

Db 992 -----AGTTGATCTTTTGT 1008
Qy 1741 GCTACCCCAAGCAATATTCAGACAGACTCATTTGAAGTTTACGAAGTTGACCTTCTT 1800
Db 1009 GCTACCCCAAGCAATATTCAGACAGACTCATTTGAAGTTTACGAAGTTGACCTTCTT 1068
Qy 1801 GTGGCTCAAGCTTTTATTTGTTAGAGACACACTTTTACAGGATGCTGATAGCAATTC 1860
Db 1069 GTGGCTCAAGCTTTTATTTGTTAGAGACACACTTTTACAGGATGCTGATAGCAATTC 1128
Qy 1861 CAGCACTTCAGAGGCTAAGGTAAATGAGTATGAGATTCCTGTGATGTTATGATACA 1920
Db 1129 CAGCACTTCAGAGGCTAAGGTAAATGAGTATGAGATTCCTGTGATGTTATGATACA 1188
Qy 1921 CGGAATTAATGGAGATAGACTTGTGCTAGAAAGGGGACTGTGTGACTGCTTATAGGC 1980
Db 1189 CGGAATTAATGGAGATAGACTTGTGCTAGAAAGGGGACTGTGTGACTGCTTATAGGC 1248
Qy 1981 AAAGTTGATGAATCCGTAATGTGCTTGAAGACGTGAGATTCAATATAGGAAT 2040
Db 1249 AAAGTTGATGAATCCGTAATGTGCTTGAAGACGTGAGATTCAATATAGGAAT 1308
Qy 2041 CCAGCTATGTGAGGTTGTTTGGAGAAATTCGATATGATGATGATGATCTCCCT 2100
Db 1309 CCAGCTATGTGAGGTTGTTTGGAGAAATTCGATATGATGATGATGATCTCCCT 1368
Qy 2101 GGACTATGCAAAATGTTGGAACCTGGTGGCAGGGGTTGCTTCTAGGTTGAGAGAC 2160
Db 1369 GGACTATGCAAAATGTTGGAACCTGGTGGCAGGGGTTGCTTCTAGGTTGAGAGAC 1428
Qy 2161 ACCAAATTAATAAATTTTAAATCTGGGACTACTATGATGATCTTATGATTTGATTA 2220
Db 1429 ACCAAATTAATAAATTTTAAATCTGGGACTACTATGATGATCTTATGATTTGATTA 1488
Qy 2221 TTGGAAGAGTGGAGTGTGAGGTTTCTCTTACCTGTGCTGCAACTATGGCAAG 2280
Db 1489 TTGGAAGAGTGGAGTGTGAGGTTTCTCTTACCTGTGCTGCAACTATGGCAAG 1548
Qy 2281 ATTGAGCCGAGCATGTGAAGACTAGTCTATGACGACACTGCAAGAAAGTTTCTTCC 2340
Db 1549 ATTGAGCCGAGCATGTGAAGACTAGTCTATGACGACACTGCAAGAAAGTTTCTTCC 1608
Qy 2341 CGCTATACAGATAGAACTCGGCTGAACCCAGAGATGTGCAAGACAGTGTTAAGTGA 2400
Db 1609 CGCTATACAGATAGAACTCGGCTGAACCCAGAGATGTGCAAGACAGTGTTAAGTGA 1668
Qy 2401 GATCTGTGTGTAACAATGATAGGCGGTGATGTGAGGCTGCTTATTTGCAAGAGCT 2460
Db 1669 GATCTGTGTGTAACAATGATAGGCGGTGATGTGAGGCTGCTTATTTGCAAGAGCT 1728
Qy 2461 GTAAGACCTCTGAAAACTTTGAAACTAATGATTAATGCAATTCGAGCTGGGCTCAAG 2520
Db 1729 GTAAGACCTCTGAAAACTTTGAAACTAATGATTAATGCAATTCGAGCTGGGCTCAAG 1788
Qy 2521 AGTAGCGTTGATGAACCTACTGTTGAATGTCCTGCTGATATGTTAAAGAGGCAAGT 2580
Db 1789 AGTAGCGTTGATGAACCTACTGTTGAATGTCCTGCTGATATGTTAAAGAGGCAAGT 1848
Qy 2581 GTGAAGATCCAGTGTGCTGCTGAGGCAATGAGACTGATTTCACTGTTCAAGCGAAGTAT 2640
Db 1849 GTGAAGATCCAGTGTGCTGCTGAGGCAATGAGACTGATTTCACTGTTCAAGCGAAGTAT 1908
Qy 2641 TTTCTTAAAGACAGCTCATCTTTCAACGCAAGGATATGTTCTTATGAAATCGAT 2700
Db 1909 TTTCTTAAAGACAGCTCATCTTTCAACGCAAGGATATGTTCTTATGAAATCGAT 1968
Qy 2701 GTGCTAACATAGGATATTAATGATGCAATTTTCAATATCTGCAATGCTCAAAATA 2760
Db 1969 GTGCTAAC----- 1978
Qy 2761 TGCTTGTTTGTGAGCTAAGAACTATGTTCCACTTAATACATGCCCAAAAGTTGTA 2820
Db 1979 ----- 1978


```
Db 133167 TCCGCAATATGAAGCTCCGACAGGTGTCATAGAAATATATGCAATGTTTTTAT 133226
Oy 1671 TTTCTTAGCATGATTAATCACTTAGGTTCTCATTTTAATGAT----- 1713
Db 133227 TTAATTTCTTCCGCCCCCTTAGTTTGTGTTATGAGGTTTACATTTATCTCTTCA 133286
Oy 1714 -----GTTGTGTAGGTGATCTTTTGTAGTACCCCAAG 1751
Db 133287 TTAATTTCTTTATCTCCCACTTTTTTTAAAGTCGAACTTTTGTAGCAACCTAG 133346
Oy 1752 CAATATTCAGACAGATCACTTTGAAGTTTACGAAGTTGACCTGCTCTTGCGCTCAAGC 1811
Db 133347 CAATATTCAGACGTAAGTTTGTAGGCCCTATGCGGTGCACTTGATGGTTCACAAAGC 133406
Oy 1812 TTTTATGTAGTAAGCAACCACTTTTACAGAGTGTGATAGCAATTCACAGCACTTCA 1871
Db 133407 CTTTGTAGTTAAAGCCCACTTTATCCAGATGCTGATATTTATTCATCAACTTCA 133466
Oy 1872 GCAGGCTAAGTAATGCTATGAGAGATTCCTGCGATGTTGATGATAC-----ACGGA 1925
Db 133467 ACAACTAAGTAACAAATATAGAGAAATGCTCCCTGTTTATCTCCATGAGATGGA 133526
Oy 1926 TTAATGGAAGATAGACTTGCTGCTTGAAGGGGACTTGTGACCTGCTTATAGCCAAAT 1985
Db 133527 GAAGAGAGAAAGTTGATTTGCAATTGAAAAGGGGCTGTGTGCACTGCTTGTGGGAGCT 133586
Oy 1986 TGATGAATGCCGTATGTGCTGAGCTTAGACAGTAGAGATTCACATATAGGAATCCAGC 2045
Db 133587 TGATCAATGTGATCATGTGTTGGTCTCGATAGTGACAGCTCACTTATCGAAACCTATC 133646
Oy 2046 TATTGTGAGTTTGTTTTGGAGAAATTCAAATCGTATGACAAATGATGATCTCCTGACT 2105
Db 133647 TATTATAGCTTTTATATGAAAACGCAAAAGGTATGAAAGCACTGATCTTCGACT 133706
Oy 2106 ATGCAATTTGTGGAACCTGCTTGGCAGGGGTGCTTCTTCTAGATTCAAGACACCA 2165
Db 133707 CTGTAAATTTGTGAGACATGGTGTATGAGAGGTGTTTCCCTAGGTTTGAATACATA 133766
Oy 2166 AGATAAAAAATTT 2178
Db 133767 AGAGCAAACTTT 133779
```

RESULT 8
AP008208_011/c

WPCOMMENT

Sequence split into 360 fragments LOCUS AP008208 Accession AP008208

Fragment Name	Begin	End
AP008208_000	1	110000
AP008208_001	100001	210000
AP008208_002	200001	310000
AP008208_003	300001	410000
AP008208_004	400001	510000
AP008208_005	500001	610000
AP008208_006	600001	710000
AP008208_007	700001	810000
AP008208_008	800001	910000
AP008208_009	900001	1010000
AP008208_010	1000001	1110000
AP008208_011	1100001	1210000
AP008208_012	1200001	1310000
AP008208_013	1300001	1410000
AP008208_014	1400001	1510000
AP008208_015	1500001	1610000
AP008208_016	1600001	1710000
AP008208_017	1700001	1810000
AP008208_018	1800001	1910000
AP008208_019	1900001	2010000
AP008208_020	2000001	2110000
AP008208_021	2100001	2210000
AP008208_022	2200001	2310000
AP008208_023	2300001	2410000
AP008208_024	2400001	2510000

AP008208_025	2500001	2610000
AP008208_026	2600001	2710000
AP008208_027	2700001	2810000
AP008208_028	2800001	2910000
AP008208_029	2900001	3010000
AP008208_030	3000001	3110000
AP008208_031	3100001	3210000
AP008208_032	3200001	3310000
AP008208_033	3300001	3410000
AP008208_034	3400001	3510000
AP008208_035	3500001	3610000
AP008208_036	3600001	3710000
AP008208_037	3700001	3810000
AP008208_038	3800001	3910000
AP008208_039	3900001	4010000
AP008208_040	4000001	4110000
AP008208_041	4100001	4210000
AP008208_042	4200001	4310000
AP008208_043	4300001	4410000
AP008208_044	4400001	4510000
AP008208_045	4500001	4610000
AP008208_046	4600001	4710000
AP008208_047	4700001	4810000
AP008208_048	4800001	4910000
AP008208_049	4900001	5010000
AP008208_050	5000001	5110000
AP008208_051	5100001	5210000
AP008208_052	5200001	5310000
AP008208_053	5300001	5410000
AP008208_054	5400001	5510000
AP008208_055	5500001	5610000
AP008208_056	5600001	5710000
AP008208_057	5700001	5810000
AP008208_058	5800001	5910000
AP008208_059	5900001	6010000
AP008208_060	6000001	6110000
AP008208_061	6100001	6210000
AP008208_062	6200001	6310000
AP008208_063	6300001	6410000
AP008208_064	6400001	6510000
AP008208_065	6500001	6610000
AP008208_066	6600001	6710000
AP008208_067	6700001	6810000
AP008208_068	6800001	6910000
AP008208_069	6900001	7010000
AP008208_070	7000001	7110000
AP008208_071	7100001	7210000
AP008208_072	7200001	7310000
AP008208_073	7300001	7410000
AP008208_074	7400001	7510000
AP008208_075	7500001	7610000
AP008208_076	7600001	7710000
AP008208_077	7700001	7810000
AP008208_078	7800001	7910000
AP008208_079	7900001	8010000
AP008208_080	8000001	8110000
AP008208_081	8100001	8210000
AP008208_082	8200001	8310000
AP008208_083	8300001	8410000
AP008208_084	8400001	8510000
AP008208_085	8500001	8610000
AP008208_086	8600001	8710000
AP008208_087	8700001	8810000
AP008208_088	8800001	8910000
AP008208_089	8900001	9010000
AP008208_090	9000001	9110000
AP008208_091	9100001	9210000
AP008208_092	9200001	9310000
AP008208_093	9300001	9410000
AP008208_094	9400001	9510000
AP008208_095	9500001	9610000
AP008208_096	9600001	9710000
AP008208_097	9700001	9810000

AP008208_098 9800001 9910000
 AP008208_099 9900001 10010000
 AP008208_100 1000001 10110000
 AP008208_101 1010001 10210000
 AP008208_102 1020001 10310000
 AP008208_103 1030001 10410000
 AP008208_104 1040001 10510000
 AP008208_105 1050001 10610000
 AP008208_106 1060001 10710000
 AP008208_107 1070001 10810000
 AP008208_108 1080001 10910000
 AP008208_109 1090001 11010000
 AP008208_110 1100001 11110000
 AP008208_111 1110001 11210000
 AP008208_112 1120001 11310000
 AP008208_113 1130001 11410000
 AP008208_114 1140001 11510000
 AP008208_115 1150001 11610000
 AP008208_116 1160001 11710000
 AP008208_117 1170001 11810000
 AP008208_118 1180001 11910000
 AP008208_119 1190001 12010000
 AP008208_120 1200001 12110000
 AP008208_121 1210001 12210000
 AP008208_122 1220001 12310000
 AP008208_123 1230001 12410000
 AP008208_124 1240001 12510000
 AP008208_125 1250001 12610000
 AP008208_126 1260001 12710000
 AP008208_127 1270001 12810000
 AP008208_128 1280001 12910000
 AP008208_129 1290001 13010000
 AP008208_130 1300001 13110000
 AP008208_131 1310001 13210000
 AP008208_132 1320001 13310000
 AP008208_133 1330001 13410000
 AP008208_134 1340001 13510000
 AP008208_135 1350001 13610000
 AP008208_136 1360001 13710000
 AP008208_137 1370001 13810000
 AP008208_138 1380001 13910000
 AP008208_139 1390001 14010000
 AP008208_140 1400001 14110000
 AP008208_141 1410001 14210000
 AP008208_142 1420001 14310000
 AP008208_143 1430001 14410000
 AP008208_144 1440001 14510000
 AP008208_145 1450001 14610000
 AP008208_146 1460001 14710000
 AP008208_147 1470001 14810000
 AP008208_148 1480001 14910000
 AP008208_149 1490001 15010000
 AP008208_150 1500001 15110000
 AP008208_151 1510001 15210000
 AP008208_152 1520001 15310000
 AP008208_153 1530001 15410000
 AP008208_154 1540001 15510000
 AP008208_155 1550001 15610000
 AP008208_156 1560001 15710000
 AP008208_157 1570001 15810000
 AP008208_158 1580001 15910000
 AP008208_159 1590001 16010000
 AP008208_160 1600001 16110000
 AP008208_161 1610001 16210000
 AP008208_162 1620001 16310000
 AP008208_163 1630001 16410000
 AP008208_164 1640001 16510000
 AP008208_165 1650001 16610000
 AP008208_166 1660001 16710000
 AP008208_167 1670001 16810000
 AP008208_168 1680001 16910000
 AP008208_169 1690001 17010000
 AP008208_170 1700001 17110000

AP008208_171 1710001 17210000
 AP008208_172 1720001 17310000
 AP008208_173 1730001 17410000
 AP008208_174 1740001 17510000
 AP008208_175 1750001 17610000
 AP008208_176 1760001 17710000
 AP008208_177 1770001 17810000
 AP008208_178 1780001 17910000
 AP008208_179 1790001 18010000
 AP008208_180 1800001 18110000
 AP008208_181 1810001 18210000
 AP008208_182 1820001 18310000
 AP008208_183 1830001 18410000
 AP008208_184 1840001 18510000
 AP008208_185 1850001 18610000
 AP008208_186 1860001 18710000
 AP008208_187 1870001 18810000
 AP008208_188 1880001 18910000
 AP008208_189 1890001 19010000
 AP008208_190 1900001 19110000
 AP008208_191 1910001 19210000
 AP008208_192 1920001 19310000
 AP008208_193 1930001 19410000
 AP008208_194 1940001 19510000
 AP008208_195 1950001 19610000
 AP008208_196 1960001 19710000

Query Match 8.7%: Score 317.6; DB 15; Length 110000;

Best Local Similarity 49.3%; Pred. No. 1.5e-60; Matches 1459; Conservative 0; Mismatches 1299; Indels 202; Gaps 16;

QY 744 TGATTTCTACAGGATTTAGAGAGCTCAACATTTCTTAACCGATGAGATCAGAGAGC 803
 Db 81995 TCGCTTGTCTGATGTTTAGGGGAGAGGCACATTCTCTGGCGATGGCATCAGAGGGC 81936
 QY 804 ATTGCAAGCTAGGAGTTTGAACCGCCGCAATTCGTTTCAGCGAGAGCTTTATCAG 863
 Db 81935 GTTCGAGCAGGATGAGCAAGCCAGCCAGATGATGATCAGCAGAGATGCTTGTGTGG 81876
 QY 864 CCGGAGCAGATTTCTTAAGCTGCTTGGCAACTGTCATATCTCGGTCTAGAAGAGA 923
 Db 81875 TCGTGAAGATGCTGCGATGATGAGCACTTCTCATGAAACGAACTCCGCACTCA 81816
 QY 924 GTACATGAGGCTCTTCTGATGATGAGAAAGCTACAGTCACTCATGATGTTCTTGGGA 983
 Db 81815 GTATGATGCTGCTTCTGAGAAACCGTGAAGAGCTCTCAGCAAGATATGCTTGGGA 81756
 QY 984 TAAAGTAATTCGATTTGGAATTAATAAGTTCT-----TCGT 1022
 Db 81755 CAAGGTGAGGCTTTAGGGGGAGCAAGGGTGTGTTGATCTTGTCAATAGACACTAGT 81696
 QY 1023 TTTATTTTCATGATTTGATTAAGAGAGAACTTTATCTAGTGAAGGTTCTCGGGCTC 1082
 Db 81695 GTATATCTCTCTGCTTAT 81636
 QY 1083 TCTGTATTCAGAGAGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1142
 Db 81635 TGTGTGGGCTTCAGAGAGGCTGGGAGGAGCACTTGTGCTGCTTGTACTGAGAGAGTTC 81576
 QY 1143 TTAAGAGAGGTTGCTTAAGTCTTTAAGCAAGATGCTTTTAACTTAAGCGCTTGGCT 1202
 Db 81575 TTTCTGATGCTCGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 81516
 QY 1203 TTTCTGATGCTCGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1262
 Db 81515 ATGTGATCTATCAAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 81456
 QY 1263 AGTTGTGAGAGAGCTTTGAAGCTTTTACAG-----TAG 1298
 Db 81455 AGGTCTCGAGAGGCTCTCAAGCTCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 81396
 QY 1299 TTTGACTTGTGTTGATTTTGAAGAGGCTTGTATTAAGAACTTCTGATTTG--- 1355

Db 81395 TTTTCATTGGGCTTAATTTGAATAGATAAATTAATTAACCTGTGAAGATTGTTT 81336
Qy 1356 -----ATACTTGTATTGAGTCTTG----- 1376
Db 81335 GGGACACATGGGCTTAATTTTATTAATAAGATGGTGAATTCCTTGAAGCAATATTAA 81276
Qy 1377 -----TGTAGAGAGAGAGACATGATGCTTGACCCGATTTAGGTGACAAA 1424
Db 81275 GGAACCTTAATCTGTAGAGAGATGAGACAGCAATCTGCACCTGATCTGCTTTCACAGA 81216
Qy 1425 TTGATAGACTTTGGAGAGATCACCCGCTTATGTCTGGAGCTTACCTGGCTTAACCG 1484
Db 81215 TTGATGAACCTTCAGAGAGATTAACCTCGCTGTATTGAGGCTTCTCTCCCTTCTTA 81156
Qy 1485 TTGTGATGATTAACCTGCGAAGAACTAAATGTTTAAGCCGTGTGCGAATATTGTT 1544
Db 81155 TTGACACAGAGCATATAAGAAAGCCCAAGAGGCTTCAAGGTGCGAAGAAACATTTGTT 81096
Qy 1545 GGTCTGTTGAGAGAGGTGAGCATCAGCTCTTGTGGGGGTTGACCCGTGAGAACTTTA 1604
Db 81095 GGAGGCTTGGCAGAGAGAGGTATGCTACCGCTTGGAGAGGATTTTCTGTGAAGCCTTCA 81036
Qy 1605 TGAATGAGGCGTTTAAAGATGACAGCTGCGAGAGGT-----ATACAGTTTAA 1655
Db 81035 TGAACGAGGCTTTTGAAGATGACATCAATTAAGAAAGGCTTCGATGTTGTTTAT 80976
Qy 1656 ATACCTTTTAAATTTCTTTAGCATGATATACTTTAGGTTTCTCATTTTAATGATGT 1715
Db 80975 GAAAGAGCTGCTATTTATTTAGCAGAGACATCTTTGAAACATGTTAACTCCCTTTC 80916
Qy 1716 TGT-----GTGTAGCTGATCTTTTGTAGCTACCCCAAGCAATATTCAGCAGATCA 1770
Db 80915 TGTTCACCTCCCAATGATGATTTCTTTCAAAAACCGAATAGACTTCCTCGTGAAGTG 80856
Qy 1771 TTGGAAGTTTACAGAGTTGCACTTGCTCTGTGCGTCAAGCTTTTATTTGTAAGAGCA 1830
Db 80855 TTGGAATTTTACATGTACACTTTGCAATGTGCGCTCAAGCAATTAATGAATGAAGCCA 80796
Qy 1831 CACCTTTTACAGAGTCTGATTAAGCAATTCAGCACTTCAGCAGCTAAAGTTAGCT 1890
Db 80795 CAATTCATCATGATGGCGGATGATCTTTTGAACAACCTCGAAGTTCAACATAGGTTCT 80736
Qy 1891 ATGGAGATTTCTGCGATGTTGATATACACGAAATTAATGGGAATGACTTCGCTCTA 1950
Db 80735 CATTAATGCTTATG-----ATATGAGATGACCTTCGATTTG 80700
Qy 1951 GAAAGGGGACTGTGCACTGCTTAATAGGCANAAGTTGAATGCGATGTGTTGGGC 2010
Db 80699 GAAAGGGCAATTTCTGCTCATTTGCTAGTCGAGATGTTAGCAAGTGAATGTGGCTTGA 80640
Qy 2011 TTAGAAGTGAAGATTCACAATATAGGAATCCAGCCTATTGTGAGTTTGTTTGAGAT 2070
Db 80639 ATGTAATATGAGTCTTCAACATACAGAGACCCCAAAATTTCTAGATTTTATGTGACAA 80580
Qy 2071 TC--AAATCGTATGACATGATGATCTCCCTGGAATATGCAAAATGTTGAAACCTGG 2127
Db 80579 TCTACATCACTGAGAGAGATGATCTTCTTCCAGGCGTGTGCACACTTTTGAGAGCTTGG 80520
Qy 2128 TTGGAGAGGGGTGCTCTTCTAGTGTTCAGAGACCAAAAGATAAAATTTAAACTGGGG 2187
Db 80519 CTTATCTTTAGAGTTTCTCTAGAGAGAGATPACTGGGGGATGCAAGTTCAAGACTTGA 80460
Qy 2188 GACTACTATGATGATCTTAATGTTTGAATTAATCTTGAAGAGTGAAGTATGATCAGGAT 2247
Db 80459 GATTACTAGATGATCCAGAAATTTTAACTTACCTTAAGAAAGATGAGAGGTGTGTGCT 80400
Qy 2248 TCTCCTTAGCTGCTGCTGCAATATGCGAAGATTTGAGC-----C 2289
Db 80399 TCTCATTTGGCTGCTGCTGCTCTATGCGAAAACCTTGGGTCTCAAGCTACAGCTCACTT 80340
Qy 2290 GAGCAGATGGAAGCTAGTGTCTATGAGAGGACATGCGAAGAAAGTTTCTTCCCGCTATACA 2349
Db 80339 GGTACTGTGAATTCAAATGCTATTCAGAGGTTCAACAGGTTTTTCTCATGTATAGAACAG 80280

Qy 2350 GATGAAGACTCGGCTGAACCCAGAGATGTGCAGAGACAGTGTTTAGTATGATCTGTT 2409
Db 80279 TTAGACAGGTGACAGCATGAGAAAATCTAAAGATGCGCTGGGGATATCTTGAAGAAATTTT 80220
Qy 2410 GGTAAATATGAGGCCCGTGAATGTGAGCCCTGG-----TGCTTTATTT 2451
Db 80219 GACCAAGAAAATGACCTGCTCATGATTTGAGAAATGCCCTTGAAGATTAATCTGTGCT 80160
Qy 2452 GCAGAACCTGTAAGACCTCTGAAAACCTTGAACCTAATGATTAATGCAATTCAGAGCTGG 2511
Db 80159 GGCGCACTGTTGACATGTTGGCAGTAAATTTGGGGCCAAATATTTGCCCTCGTAAGAGCCC 80100
Qy 2512 GTCTCAGAGATGACGTTGATGAATCTAATGTTGAAATGTCCGTTGCTGATATGTTAAAG 2571
Db 80099 CTTTCTGCTATATGAGATGAGATGATGATCTGTGGCAGTTGCTAAATGATGTGCACTACT 80040
Qy 2572 GAGCAGATGTAAGATCTTACGTCTGTGTGTGCAATGGAATGATTTCACTGATTTCACTTACG 2631
Db 80039 GATGTGATGAATGCTTTTCATT--TGCTAAAGTATGTCAAACACTGGAATTTACTG 79982
Qy 2632 CAGAAATATTTCTTAAGAGAGCTCATCTTTTCAACGCAAGATATGTTTCTTATG 2691
Db 79981 CTACGCTGTTTACTGGGATTCCTTGTGGCTGATTAAGATATGTTCTTGATCTTATAGCT 79922
Qy 2692 GAATCTGATGTGCTACCATAGTATGATTAATGATGCAATTTTCATATATCTGACTG 2751
Db 79921 TTGCATTAATGTTTAAACAAAGCTATACATCAAAATTTTGTGCTCATACATACACATG 79862
Qy 2752 CTCAAAATATGCTGTTTGTGTGAGCTAAGAACATAGTCCCACTTAATATATGTCCTCAA 2811
Db 79861 CTACTAATATCTGTAATGTAATATCATAGAGATTAATTTTGATGATGATATGTTAAG 79802
Qy 2812 AGTGTACCAAGATTAACAAGTGTGCTGATGAATTTCACTAATATGCTGCTTG-----A 2866
Db 79801 ATCTGCACTAGATGAAGATTCAGTACATATTCCTGAATGAGATGCCAAGCTGGCAGAA 79742
Qy 2867 ATTTTGTGATCAACGTGTAGACAGAAATGTAATTTCACTCTCAACTTTCTGTTTGA 2926
Db 79741 ATATTTGTCGCAAGTGTGACAGATGATCAAAATCTAAGGCTTGGGACCAAACTTCGGTTG 79682
Qy 2927 TAACTGAGATTAAGATTCCTTATGTTGCTGCTTGTGCCAATTTTTC--CTGTATTT 2985
Db 79681 CATCATTTGCAAGAGTATGTTTAAACTTTGACTTTTCAATGCTTCTTACGTGGCCACTG 79622
Qy 2986 TTTTCTTTTGCATTTAAGGTCAGTCAGAGCTGACGATTTCAAGAGCACTTCCAGAAATGA 3045
Db 79621 CTCTTACCTATACCTTACGCTTCTTATAGAGATTAATTCAGATTAATCTCTCTATAGGA 79562
Qy 3046 TGTAGGACTGCAAGAAATATAGTATCCAAATGTCAGAGAAATTAAGTCTGTGCTTTTGG 3105
Db 79561 TATATGCTACGCAATGTACTTTTCAAGTTTGTGTTTCAATATCTTTCT----- 79511
Qy 3106 GCTGTATCACCGCATTAAGAAATGTTTACAGAGTGAAGGAATTAATCTACAAATTCATCA 3165
Db 79510 ---TGTTTGAATCAACACATGCTTCAATGCACTGAACAAATTTTCAAGAACTATATGC 79454
Qy 3166 TTGTGTAAAACTGTGGAATGATTAATGATGCTGCTGCTTTGTAAT---CTGTATTT 3222
Db 79453 CACTGCAATGATATCCCTGATGATGAGCGGTCAAAATGCTGTACTGATTTTCCATGTCT 79394
Qy 3283 AGCTTGGGTGTTTATGATTAATCACTGTGTTGAACATATCTGTGTAACATGTGCACTCT 3342
Db 79333 GTCAATGGGTGTTTGTGGAATTAATCACTATTCGATGGAAGATTAAGATCACTATCT 79274
Qy 3343 CAGCAGATGGAACCCGTCTCTGTGTGAAGCACTCTGAGAGATGCTGTCTATCTG 3402
Db 79273 CCTATGATGTGCAAGAGCACTGTGAGAGGTACGATTAATGAGGAGGCCAACTTACTG 79214

Oy 3403 ATTGGTTCATCCAGAAAAGTCTAGTATGTCAGAACTACACAGACAGATCGAAG 3462
 Db 79213 ATGTTACTAGGCCCAAGAAACATGATTCATATGACCAAAATATACCTACCCGTAATGACA 79154
 Oy 3463 TTTTCTGCGCCAG--TAAAGGTGAAATAATCATGAGAGCTCTGTTCTTGATCATATAT 3519
 Db 79153 TGGCCTCTCCAGAGCTAGAGGAGTGAAGATTAACGAAAGACAGCTCTCAAGTCGTAGA 79094
 Oy 3520 ATACTCATATGTCAGATGTC 3539
 Db 79093 ACGAGTGTGACATGACTTGTTC 79074
 RESULT 9
 LOCUS AP004885/c 150462 bp DNA linear PLN 15-SRP-2004
 DEFINITION Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 2,
 PAC clone:P0575F10.
 ACCESSION AP004885
 VERSION AP004885.3 GI:41053009
 SOURCE
 ORGANISM
 Oryza sativa (japonica cultivar-group)
 Oryza sativa (japonica cultivar-group)
 Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzeae; Oryza.
 REFERENCE
 AUTHORS 1 Sasaki, T., Matsumoto, T. and Yamamoto, K.
 TITLE Oryza sativa nipponbare (GA3) genomic DNA, chromosome 2, PAC
 clone:P0575F10
 JOURNAL Published Only in Database (2002)
 REFERENCE 2 (bases 1 to 150462)
 AUTHORS Sasaki, T., Matsumoto, T. and Yamamoto, K.
 TITLE Direct Submission
 JOURNAL Submitted (20-MAR-2002) Takuji Sasaki, National Institute of
 Agricultural Sciences, Rice Genome Research Program, Kannondai
 2-1-2, Tsukuba, Ibaraki 305-8602, Japan
 (E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/
 Tel:81-298-38-7441, Fax:81-298-38-7468)
 On Jan 21, 2004 this sequence version replaced gi:38142429.
 Genes were predicted from the integrated results of the following:
 GENSCAN (http://CCR-081.mit.edu/GENSCAN.html), FGENESH
 (http://www.softberry.com/), GeneMark.hmm
 (http://opal.biology.gatech.edu/GeneMark/), GlimmerM
 (http://www.tigr.org/db/glimmer/glmr_form.html), RiceHMM
 (http://rgp.dna.affrc.go.jp/RiceHMM/), SplICEPredictor
 (http://bioinformatics.iastate.edu/cgi-bin/sp.cgi), sim4
 (http://www.tigr.org/software/glimmer/), BLASTN and BLASTX. The
 genomic sequence was searched against NCBI Nonredundant protein
 database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA
 sequence database at RGP or DBJ. Protein homologues of the coding
 regions were searched against NCBI Nonredundant protein database
 with BLASTP. ESTs represent the identified cDNA sequences using
 BLASTN with the corresponding DBJ accession no. and RGP clone ID.
 Full-length cDNAs represent the identified cDNA sequences using
 BLASTN with the corresponding DBJ accession no.
 A gene with identity or significant homology to a protein is
 classified based on the protein name to indicate the homology level
 such as same name, 'putative-' and '-like protein'. A gene without
 significant homology to any protein but with full-length cDNA or
 EST homology (covering almost the entire length of partial
 sequence) is classified as an 'unknown' protein. A gene predicted
 by two or more gene prediction programs is classified as a
 'hypothetical' protein according to IRGSP standard. A gene
 predicted by a single gene prediction program is also classified as a
 miscellaneous feature of the sequence.
 The orientation of the sequence is from T7 to SP6 of the PAC clone.
 This sequence of P0575F10 clone has an overlap with OJ1020 C02
 (DBJ: AP004078) clone at 5' end and with P0482P12 (DBJ: AP005111)
 clone at 3' end. Detailed information on overlap and assembly
 quality together with annotation of this entry is available at

http://rgp.dna.affrc.go.jp/GenomesSeq.html.
 FEATURES
 source
 Location/Qualifiers
 1..150462
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="genomic DNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /chromosome="2"
 /clone="P0575F10"
 complement(4987..8112)
 /gene="P0575F10.1"
 complement(join(4987..5412,5520..5679,5764..5914,
 6096..6247,6738..6859,7864..8112))
 /gene="P0575F10.1"
 /note="supported by full-length cDNA(s): AK106553"
 complement(join(5003..5412,5520..5679,5764..5914,
 6096..6247,6738..6859,7864..8084))
 /gene="P0575F10.1"
 /note="supported by full-length cDNA(s): AK073059"
 complement(join(5315..5412,5520..5679,5764..5914,
 6096..6247,6738..6859,7864..7876))
 /gene="P0575F10.1"
 /note="contains EST(s): AU070160(BE1534)
 contains full-length cDNA(s): AK106553,AK073059"
 /codon_start=1
 /product="putative eukaryotic translation initiation
 factor 3, subunit 1 alpha"
 /protein_id="BAD07942.1"
 /db_xref="GI:41053010"
 /translation="MEDWDSDFQPAVPSAKAPLKSMAEDVEDDEVDKSWEEER
 EKPRPPEVETKAPKPSGAKKQQAQSTSEVQDADDPALKEKRLVEAD
 FSKTTELPFGKDGSEKSLDTPIPKSESPFAEYAIANKLPYEKSPKMLKVMR
 LSMALKGADKDISSTIAIANKIKAEKKAAGKKQAKKQLHIENKDDPIPG
 KGNFDDPDREVDPM"
 complement(10186..13754)
 /gene="P0575F10.2"
 complement(join(10186..10479,10759..10892,11127..11991,
 12089..12349,12515..12712,12843..13073,13455..13754))
 /gene="P0575F10.2"
 /note="start and end point are not identified"
 complement(join(10186..10479,10759..10892,11127..11991,
 12089..12349,12515..12712,12843..13073,13455..13754))
 /gene="P0575F10.2"
 /note="contains EST(s): AU058423(B51136), AU183658(B51136)"
 /codon_start=1
 /product="lactid division protein precursor"
 /protein_id="BAD07942.1"
 /db_xref="GI:41053011"
 /translation="MEGFHNLARPNNSAPFSLPRPRRRPPPHPSACRAAR
 WAERLFADPHLPRAESDPPSPAPAPAAASAPFVLPFDAAERSLPLQVDFKVL
 GAEHPLGDDGIRRAFEARIAKPPQYGTSDALVGRQMLQIAHPTLMQNSRTQYDRA
 LSKREERLMDIMDKBEGALAVIVATGEOILLDPKRRKQVVLAMALAYDLR
 DNAAAPEDVAGCEVLERALKLQEOGASLADLDLSQIDETLEIRPRVLELISL
 FIDTEHKKRQEGVAGANILMSVGRGIALVVGGSFEAPMEAFVWTSIEQDF
 SKTPNSLPPEFELYNALAHVAQALISKRPQTMADLEQLQKFNIGSHVADNS
 MDLLEAFACSLVLDGSKCMWGLINSSPPYDPKTLERIVNNSISSENDLPLGL
 CKLETWLIFVPPRSRDTKQMPRLGAYDDPYVLSTLEMEGGAHLAAAAIAK
 LGAATATAITGVKSNATQAEKVPRLLEQIDRMENTKQDGGVLENFDEENPAHD
 SRNALKISKALPALAVYGATLPRKRLSIRBSGVAIVANSVDSDDPRLDDE
 DPVHPNDADIEDIVKWKOSISKALGESHVASLOEVDGNLKWTRAEIER
 HGWWEYTLSDVTLDISITISIDGRRAVEATIDEAGQLDVTBPRNDSYDKTTRY
 EMASFSGKGMKTEGAVLKS"
 complement(16401..17341)
 /gene="P0575F10.3"
 complement(join(16401..16739,16841..17341))
 /gene="P0575F10.3"
 /note="start and end point are not identified"
 complement(join(16401..16739,16841..17341))
 /gene="P0575F10.3"
 /note="similar to Arabidopsis thaliana chromosome 3,
 At3g61870"
 /codon_start=1
 /product="hypothetical protein"

CDS

mRNA

gene

/protein_id="BAD07943.1"
/db_xref="GI:41053012"
/translation="MMAAPLAAVHAIINCASNNKSPSPAROOQOTTTTATRGSPAA
LPSLRTTAAAAATATLALAPDPAALGEGGIIIEGRVALHPVWGGLAATYLMVEI
YLGOMRRVRTIODEINELKKQKRAAAAATPAVAADSSSSSSPPSPAPKPEVEI
KIDLETERRKLIIGSPFRDRHFNAGSIIIGGVTESSVGAALNTWFRKGLPQPHLEA
GPAITVMAAAAAALVPAMQKNETARSLHIALMINVLLPTWQIPTEIIVGKEEFT
TTP"
complement(18499..18942)
/gene="P0575F10.4"
complement(<18499..>18942)
/gene="P0575F10.4"
/note="start and end point are not identified"
complement(18499..18942)
/gene="P0575F10.4"
/note="contains EST (s): AU084851(R10374), AU084850(R10374)"
/codon_start=1
/product="putative EF-hand Ca2+-binding protein CCD1"
/protein_id="BAD07944.1"
/db_xref="GI:41053013"
/translation="NNHHQIIRSTTAARQESAGGGGGEEDLMPVAGRLGAEGL
LSELRAGRLADPARGAITBASLRSAASYIGGGGGEMTVEEAAAMVREGQDDG
GALSEARCVTMVRLSPGIMDABGMLAEALDELRLPLPPPPA"
21937..23552
/gene="P0575F10.5"
join(<21937..22225,22326..22648,22736..23049,
23393..>23552)
/gene="P0575F10.5"
/note="start and end point are not identified"
join(21937..22225,22326..22648,22736..23049,23393..23552)
/gene="P0575F10.5"
/note="similar to Arabidopsis thaliana chromosome 5,
At5g24330"
/codon_start=1
/product="hypothetical protein"
/protein_id="BAD07945.1"
/db_xref="GI:41053014"
/translation="MGAPTLPLRRTRPAATRAEGSGGDGDDDVRCACGSGESA
AELLDCDCDGLHIFCLRPILRPVADWFCPSGASPSKSKHAKKQCPVQ
KIDVDFIORGPAALAAASSEKGRKRVGIRLVSKKKRLPFPNSDPAR
LRQMSIATLALGACVAFSNELTYPGMAPAARPAALAESEGMOLPREDENTICK
RMAARGEPPLIVYDPEGFTVEADRIKLTITTEYGVGDVLTREHHDGSMWT
LNDISATPSRSLVLCPKRSNIARFLNGINHTPDGRKQNLKCVRFDVGSGCRVLVA
NRISKSGRLYYDNGSEHEYPTHFV"
complement(24479..24817,24913..25266,25398..25633,
25720..25832,27054..27185,27995..28148)
/gene="P0575F10.6-1"
complement(join(24479..24817,24913..25266,25398..25633,
25720..25832,27054..27185,27995..28148))
/gene="P0575F10.6-1"
complement(join(24479..24817,24913..25266,25398..25633,
25720..25832,27054..27185,27995..28148))
/note="supported by full-length cDNA (s): AK121372"
complement(join(24491..24817,24913..25266,25398..25633,
25720..25832,27054..27185,27995..28147))
/gene="P0575F10.6-1"
complement(24563..28147)
/gene="P0575F10.6-2"
complement(24563..28147)
/gene="P0575F10.6-2"
/note="contains full-length cDNA (s): AK065148
non-coding transcript
probably inactive due to including stop codon(s) in CDS"
complement(join(24787..24817,24913..25266,25398..25633,
25720..25832,27054..27159))
/gene="P0575F10.6-1"
/note="contains EST (s): AU031017(E60571), AU093466(B60571)
contains full-length cDNA (s): AK121372, AK062255"
/codon_start=1
/product="putative arginine/serine-rich splicing factor

Query Match 8 7%; Score 317.6; DB 15; Length 150462;
Best local Similarity 49.3%; Pred. No. 156-60;
Matches 1459; Conservative 0; Mismatches 1299; Indels 202; Gaps 16;

QY 744 TGATTTCTACAGGATATTAGAGCTCAACACATTTCTTAACCGATGAATCAGAGAGC 803
DB 13086 TGGCTTTTGCTCAGGTTCTTAGGGGCGAGAGCCACATTTCTTGCGGATGGCATCAGAGAGGC 13027
QY 804 ATTGGAAGCTAGGGTTTTCGAAACCGCCGCAATTCGGTTTACGAGCAGACGTTTAATCAG 863
DB 13026 GTTCGAGGACCGGATACCAAGCCAGCATATAGGCTTACAGCAGCAGATGCTCTTGTGG 12967
QY 864 CCGGAGCAGATTTCTTCAAGCTGCTTGCAGAACTCTGTAAATCTCGGCTTGAAGAGA 923
DB 12966 TCGTCGACAAATGCTGCAGATTGCCATGACCTCTCATGAACCAAACTCCGCACTCA 12907
QY 924 GTACAAATGAAGCTTCTTGTGATGAATGAAGAAGCTACAGTCATCACTGATGTTCTTGGGA 983
DB 12906 GTATGATCGTGCCTTTCTGTGAACCGTGAAGAAGCTCTCACATGTGATATTGCTTGGGA 12847
QY 984 TAAGTAAATTCGATTTCCGAATATATAAGTTCTT-----TCGT 1022
DB 12846 CAAAGTGAAGCTTTAAGGGGCGAGCAAAAGGGTGTGTGATCTTGTCAAAATAGACATAGT 12787
QY 1023 TTTAATTTCAATGAATGTGATTAAGAAAGAACTTTATCTAGTGAAGTTCTGGGCTC 1082
DB 12786 GTATTAATCTCTCTGCTTATATATATTGGGCTCTTGTGAAGACAGTTCCAGGTGTG 12727
QY 1083 TCTGTGATTTGCAAGAAAGTGTGAGACTGAGATAGTTCTTCGGGTGTGAGGCTGTGC 1142
DB 12726 TGTGTGCGCTTCAAGAGGCTGAGGAGGCACTTGCTGTGCTTTGACTGGAACAGTTGC 12667
QY 1143 TTAAGAGAGGTTCCCTTAAGCTGTTTAAGCAAGATGTGCTTTTATGTTATGGCGTTCCGT 1202
DB 12666 TTTCTGATTCGGCCACCAAGGCTTCAAGAGAGCGTGGTCTAGCCATGCGCTGTGCTT 12607
QY 1203 TTTCTGATGTCTGAGGGAATGCTATGGCACTTGATCCACCTGATTTTATTACTGTTATG 1262
DB 12606 ATGTGATCTATCAAGGAATGCTATGGCAGCAAGCCCTCCAGATGTATTTGGCTGTCCG 12547
QY 1263 AGTTTGTGAAGAGCTTTGAAGCTTTTACAG-----TAG 1298
DB 12546 AGGTGCTCGAAGAGGCTCTCAAGCTTTGCAAGTGAATGTTGCTTTTATCTCCCTAC 12487
QY 1299 TTTGACTGCTTTGTGTAATTTGACAGACGTTGCTTTATAAGAACTTTCTTGATTTG--- 1355
DB 12486 TTTTTCATTTGGCTTATATTGATTAAGATTAAGATTAAATTAACCTGTGAAGATTTGTT 12427
QY 1356 -----ATACCTTTGTTATTAAGCTTG----- 1376
DB 12426 GGGACATGGGCTATATTTTTTTTATTAAGAAATGTAATTTCTTGAAGCAATTTAATA 12367
QY 1377 -----TGTAGAGAGAGCAAGTACCTTGACCGGATTTACGTGACAAA 1424
DB 12366 GGAACCTCATCTGTAGGAAGATGGAACAGCAATTCGCACTGATCTGCTTTACAGAGA 12307
QY 1425 TTTGATGACCTTTGGAAGAATCATCTCCGCGTTATGTCTTGAAGCTACTTGGCTTAACCC 1484
DB 12306 TTTGATGAACCTCTGAGAGATTTACACCTCGCTGTATTTGAGACTTCTCTCCCTCTTA 12247
QY 1485 TTTGATGATTAACGCTGCGAAAGAATAATGATTTAAGCGGTGTCGGAATTTTGT 1544
DB 12246 TTGACAGAGACATCAATTAAGAGCGCCAGAGAGGCTTCAAGGTGCGAAGAAACATTTGT 12187
QY 1545 GGTGTGTTGAGAGAGTGAAGCATCAGCTCTTGTGGGAGTTTGAACCGTGAAGATTTTA 1604
DB 12186 GGAAGCTTTGGAGAGGATTTATGCTACCTGTGAGAGGATTTTCTGTGAAGACCTTCA 12127
QY 1605 TGAATGAGGCTTTTACGATGACAGCTGTGAGCAGGT-----ATACAGTTTGA 1655
DB 12126 TGAACGAGGCTTTTGAAGATGACATCAATTGAACGGTTCGATCATGTTTGTGTTTAT 12067
QY 1656 ATACCTTTTATATTTCTTAGCATATATACTTAGGTTTCTCATTTTATGATGT 1715
DB 12066 GAAACAGCTGCTTATTTATATTGAGAGACATCTTTGAAAACATGTTAATCCTCCTTTC 12007

QY 1716 TGT-----GTGGAGTGTGATCTTTTGTGACTACCCAGCAATATTCAGAGATCA 1770
 DB 12006 TGTTCACCTCCAGATGATCTTTTCAAAAAACCGAATAGCATTCCTCCTGAATGG 11947
 QY 1771 TTGGAAGTTAGGAATTGCACTTGCTCTTGCGCTCAAGCTTTTATGTGAAGACCA 1830
 DB 11946 TTGGAATTTTACAAATGTACACTTGCACATGTCGCTCAAGCAATTAATGAAGGCCA 11887
 QY 1831 CACCTTTTACAGAAATCTGATTAAGCAATTCAGCAACTTCAGAGGCTTAAGGTAATGCT 1890
 DB 11886 CATTCATCATGATGAGCGGATGATCTTTTGAACAACCTCAAGAGTTCAACATAGTTCT 11827
 QY 1891 ATGAGATCTCTGCGATGTTGATGATACAGGAATTAATGGAGATTAAGCTTGGCTTA 1950
 DB 11826 CATTAATGCTTAATG-----ATATAGAGATGAGCCTTGATG 11791
 QY 1951 GAAAGGGGACTGCTGTCATGCTTATAGGCAAAAGTTGAATGCGGTATGAGTGGGC 2010
 DB 11790 GAAAGGGCAATCTGCTCATATGCTAGTCGAGAGATGTTAGCAAGTGCAGAAATGGCTTGA 11731
 QY 2011 TTAGACAGTGAAGATTCACATATAGGAATCCAGCTATGTTGAGTGTGTTTGGAGAT 2070
 DB 11730 ATTGAATTAAGATCTTCAACATACAGAGACCCCAAAATTTCTAGAGTTTATGTGACAAAC 11671
 QY 2071 TC---AAATGTGATGACAAATGATGATCTCCCTGACATAGCAAAATGTTGGAACCTGG 2127
 DB 11670 TCTAGCATGATGAGAGATGATCTTCCAGGGGCTGTCAGACTTTTGGAGACTTGG 11611
 QY 2128 TTGGCAGGGGTGCTCTTCTTCTAGTTTCAGAGACCAAAAGATTAATACTCGGG 2187
 DB 11610 CTATATCTTTGAGGTTTCTTGAAGACAGATATCTCGGGGCAATGCAATGCACTTGG 11551
 QY 2188 GACTATATGATGATCTTATGTTTGAATCTTCTTGAAGAAGTGAAGTGAAGTGAAGT 2247
 DB 11550 GATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 11491
 QY 2248 TCTCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2289
 DB 11490 TCTCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 11431
 QY 2290 GAGCATGATGAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2349
 DB 11430 GGTACTGATGAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 11371
 QY 2350 GATGAAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2409
 DB 11370 TTAGACAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 11311
 QY 2410 GGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2451
 DB 11310 GATCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 11251
 QY 2452 GATGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2511
 DB 11250 GGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 11191
 QY 2512 GCTCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2571
 DB 11190 CTTTCTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 11131
 QY 2572 GAGGCAAGTGAAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2631
 DB 11130 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 11073
 QY 2632 CAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2691
 DB 11072 CTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 11013
 QY 2692 GATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2751
 DB 11012 TTGCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 10953
 QY 2752 CTCAAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2811

DB 10952 CTACTATTAATCTGTAATGATATCATGAGAGTTTACCTTTGATGATGATGATGATGATGATGATGAT 10893
 QY 2812 AGTTGATCAAGATTAATCAAGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2866
 DB 10892 ATCTGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 10833
 QY 2867 ATTTTGTATCAAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2926
 DB 10832 ATATTGTTGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 10773
 QY 2927 TAAGTGAATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2985
 DB 10772 CATCATTCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 10713
 QY 2986 TTTTCTTTTCAATTAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3045
 DB 10712 CTTTACTATTAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 10653
 QY 3046 TGTGAT 3105
 DB 10652 TATATGAT 10602
 QY 3106 GCTGAT 3165
 DB 10601 ---TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 10545
 QY 3166 TTGAT 3222
 DB 10544 CACTGAT 10485
 QY 3223 TATAGAT 3282
 DB 10484 TATAGAT 10425
 QY 3283 AGCTGAT 3342
 DB 10424 GTCATGAT 10365
 QY 3343 CAGCATGAT 3402
 DB 10364 CCTGAT 10305
 QY 3403 ATTTGAT 3462
 DB 10304 ATTTGAT 10245
 QY 3463 TTTTCTGAT 3519
 DB 10244 TGGCTTCTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 10185
 QY 3520 ATATGAT 3539
 DB 10184 ACATGAT 10165

RESULT 10

AC155597

LOCUS

DEFINITION

AC155597

VERSION

KEYWORDS

SOURCE

ORGANISM

Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD clade; Panicoideae; Andropogoneae; Zea.

REFERENCE

AUTHORS

Chan, A.P., Pertea, G., Zheng, L., Cheung, F., Lee, D., Koo, H., Utecher, T.R., Feldblum, T.V., Rabinowicz, P., Fraser, C.M., Schubert, K., Samiguel, P., Ma, J., Pontaroli, A.C., Rohlfing, T.,

TITLE
JOURNAL
REFERENCE
AUTHORS
Budiman,A., Bedell,J.A., Lakey,N., Barbazuk,B., Bennetzen,J. and Quackenbush,J.
Consortium for Maize Genomics - BAC skim sequencing and assembly unpublished
2 (bases 1 to 184752)
Chan,A.P., Perlea,G., Zheng,L., Cheung,F., Lee,D., Koo,H., Uterback,T.R., Feldblum,T.V., Rabinowicz,P., Fraser,C.M., Schubert,K., Samiuel,P., Ma,J., Pontaroli,A.C., Rohlfing,T., Budiman,A., Bedell,J.A., Lakey,N., Barbazuk,B., Bennetzen,J. and Quackenbush,J.
Direct Submission
Submitted (15-JAN-2005) The Institute for Genomic Research (TIGR), 9712 Medical Center Dr, Rockville, MD 20850
3 (bases 1 to 184752)
Chan,A.P., Perlea,G., Zheng,L., Cheung,F., Lee,D., Koo,H., Uterback,T.R., Feldblum,T.V., Rabinowicz,P., Fraser,C.M., Schubert,K., Samiuel,P., Ma,J., Pontaroli,A.C., Rohlfing,T., Budiman,A., Bedell,J.A., Lakey,N., Barbazuk,B., Bennetzen,J. and Quackenbush,J.
Direct Submission
Submitted (25-JAN-2005) The Institute for Genomic Research (TIGR), 9712 Medical Center Dr, Rockville, MD 20850
On Jan 25, 2005 this sequence version replaced gi:57863118.
-----Trace Submission-----
Center name: TIGR
Seq lib id: ZGGO
-----Project Information-----
Web site: <http://www.tigr.org/tdb/tgi/maize/>
Contact: maize@tigr.org

* NOTE: This is a 'working draft' sequence. It currently consists of 26 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
* 1
* 2210: contig of 2210 bp in length
* 2311 2310: gap of unknown length
* 9942 9941: contig of 7631 bp in length
* 10042 10041: gap of unknown length
* 16745 16745: contig of 6704 bp in length
* 16746 16845: gap of unknown length
* 16846 32394: contig of 15549 bp in length
* 32395 32494: gap of unknown length
* 32495 35494: contig of 3000 bp in length
* 35495 35594: gap of unknown length
* 35595 40051: contig of 4457 bp in length
* 40052 40151: gap of unknown length
* 40152 41379: contig of 1228 bp in length
* 41380 41479: gap of unknown length
* 41480 44517: contig of 3038 bp in length
* 44518 44617: gap of unknown length
* 44618 45564: contig of 947 bp in length
* 45565 45664: gap of unknown length
* 45665 49440: contig of 3776 bp in length
* 49441 49540: gap of unknown length
* 49541 52698: contig of 3158 bp in length
* 52699 52798: gap of unknown length
* 52799 55964: contig of 3166 bp in length
* 55965 56064: gap of unknown length
* 56065 59236: contig of 3172 bp in length
* 59237 59336: gap of unknown length
* 59337 70052: contig of 10716 bp in length
* 70053 70152: gap of unknown length
* 70153 84603: contig of 14451 bp in length
* 84604 84703: gap of unknown length
* 84704 86418: contig of 1715 bp in length
* 86419 86518: gap of unknown length
* 86519 88069: contig of 1551 bp in length
* 88070 88169: gap of unknown length
* 88170 89661: contig of 1492 bp in length

FEATURES
source
* 89662 89761: gap of unknown length
* 89762 91357: contig of 1596 bp in length
* 91358 91457: gap of unknown length
* 91458 132727: contig of 41270 bp in length
* 132728 132827: gap of unknown length
* 132828 137468: contig of 4641 bp in length
* 137469 137568: gap of unknown length
* 137569 163843: contig of 26275 bp in length
* 163844 163943: gap of unknown length
* 163944 169858: contig of 5915 bp in length
* 169859 169958: gap of unknown length
* 169959 176100: contig of 6142 bp in length
* 176101 176200: gap of unknown length
* 176201 183079: contig of 6879 bp in length
* 183080 183179: gap of unknown length
* 183180 184752: contig of 1573 bp in length.
Location/Qualifiers
1. 184752
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/c1one="ZM58B0216K08"
2211. 2310
/estimated_length=unknown
9942. 10041
/estimated_length=unknown
16746. 16845
/estimated_length=unknown
32395. 32494
/estimated_length=unknown
35495. 35594
/estimated_length=unknown
40052. 40151
/estimated_length=unknown
41380. 41479
/estimated_length=unknown
44518. 44617
/estimated_length=unknown
45565. 45664
/estimated_length=unknown
49441. 49540
/estimated_length=unknown
52699. 52798
/estimated_length=unknown
55965. 56064
/estimated_length=unknown
59237. 59336
/estimated_length=unknown
70053. 70152
/estimated_length=unknown
84604. 84703
/estimated_length=unknown
86419. 86518
/estimated_length=unknown
88070. 88169
/estimated_length=unknown
89662. 89761
/estimated_length=unknown
91358. 91457
/estimated_length=unknown
132728. 132827
/estimated_length=unknown
137469. 137568
/estimated_length=unknown
163844. 163943
/estimated_length=unknown
169859. 169958
/estimated_length=unknown
176101. 176200
/estimated_length=unknown
183080. 183179
/estimated_length=unknown

ORIGIN

Query Match 6.7%; Score 246.8; DB 14; Length 184752;
 Best Local Similarity 57.0%; Pred. No. 1.3e-44;
 Matches 564; Conservative 0; Mismatches 372; Indels 54; Gaps 4;

QY 1329 TGGCTTATAGAACTTTCTTGATTTG--ATACTTGTATTGAGTCTTGTGAGAGGA 1386
 |||||
 DB 108883 TGTCTTGATTAAGATATCTTGTAAGCACTACTAATGAAACATTTTACTTTAGAGGA 108942
 |||||

QY 1387 AGGAGCAAGTACGCTTGACCGGATTTAGTCCACAATTTGATGAGACTTTGGAAGAT 1446
 |||||
 DB 108943 TGGTGAAGCAATCTTGACCTGAACTCTTGCAAGATTGAACTTTGAGAGAT 109002
 |||||

QY 1447 CACTCGCGTTATGCTTGAAGCTACTTGGCTTACCGCTGGTGAATGATTCGTCGCA 1506
 |||||
 DB 109003 TACACTCGTTGTATTTGAGACTTCTCTCTCTACTGATGAAAAATTA 109062
 |||||

QY 1507 AAGACTAAATGCTTTAAGCGGTGCGGAAATTTTGTGCTGTTGAGAGAGTGAAC 1566
 |||||
 DB 109063 ACGCAAGAAAGCTTGCAAGGTGCAAAAAACATATTGTGAGTGTGGCAGGGGTGTAT 109122
 |||||

QY 1567 ATCAGCTCTTGTGGGGGTTTGAACCGTGAAGATTGATGAGCGCTTTTACGAT 1626
 |||||
 DB 109123 TGTCTACTGTTGAGAGGATTTTCTGAGGCTTCATGAATGAGGCTTTCTTGCAGAT 109182
 |||||

QY 1627 GACAGCTGCTGACAGGT-----ATACAGTTTATGATTA 1661
 |||||

DB 109183 GACATCAGCTGAGAGGTTTGTGTCTACCAAGTTTAATGAGCAATCTGATGTGTGA 109242
 |||||

QY 1662 TTTTAAATTTCTTGAAGATGATTAATCTTAAAGTTTCTCATTTTAAATGATGTGTG 1721
 |||||

DB 109243 GTCTGTCCATTTTACTTGAAACAGATTTTAAAGCTTCACTCACTTTCTGTTAATATC 109302
 |||||

QY 1722 GTAGGTTATCTTTTGTATGATCAACCAAGCAATATTCAGAGAGTCACTTGAAGTTA 1781
 |||||

DB 109303 TCAGATGATTTCTTCTTAAACACCAATGACATCACTGATGATGTTTGAATCTA 109362
 |||||

QY 1782 CGAATGTCATCTGCTCTGTGAGTCAAGCTTTTATTTGTAAGAACCACTTTTACA 1841
 |||||

DB 109363 TAGTGTCCGACTTGGCCCACTGTCTCAAGCAATTCGAAGTAAAGGCCAATTCATCAT 109422
 |||||

QY 1842 GATGCTGATTAAGCAATTCAGCAACCTTCAGAGGCTTAAGTATGCTATGAGATTC 1901
 |||||

DB 109423 GATGCAAGTGAACCTTCTGCAACCTTCAGAACTTCAATATAGGTTCTCAATATCTTA 109482
 |||||

QY 1902 TGGCATGTTGATGATCAACGGAATATTTGGAGATGATGCTTGGCTTAAGAAAGGAGCT 1961
 |||||

DB 109483 TGAATAAC-----GAGATGACCTTCGATTAAGAAAGGAGCT 109518
 |||||

QY 1962 CTGTGCACTGCTTATAGGCAAGTATGATGATGCGTATGCTTGGGCTTGAAGATGA 2021
 |||||

DB 109519 CTGCTGCTTACTTGTGGGGGATTTTGCATTTGCAATGAGATGCTTGAATGATTAATGA 109578
 |||||

QY 2022 GATTTCAATATTAAGAAATCCAGCTATTGTGAGTTTGTTTTGAAGATTC---AAATCG 2078
 |||||

DB 109579 GTCTCTGCTATATAGGGAACCCCAAAATATCGAGTTTGTGATTAATCTTAAGATGAA 109638
 |||||

QY 2079 TGATGCAATGATGATCTCCCTGCACTATGCAAAATTTGTTGAAACCTGTTGGCAGGGGT 2138
 |||||

DB 109639 CGAAAGATATGATCTTCTTCCAGGGGCTGTGCAAGCTTTGAGACATGCTGTGTCTCCGA 109698
 |||||

QY 2139 TGTCTTCTTAGTTCAGAGACCAAAAGATTAATACTCGGGGACTACATGA 2198
 |||||

DB 109699 GGTATTTCCAGAGAGACAGATACCCGAGGAGCGAGTTTCAGCTGGGAGACTATTAAGA 109758
 |||||

QY 2199 TGATCTATGTTTGAAGTTACTTGAAGAAAGATGAGGTGATTCAGGGGTTCTCTTTAGC 2258
 |||||

DB 109759 GCAACCAAAAGTTCTTAAGCTACTTAGAAAGATGAGAGCGGTGTGCTTCCATTTAGC 109818
 |||||

QY 2259 TGTGCTGCACTATGAGAGATTTGAGC 2288
 |||||

DB 109819 TGCAGCTGCGCTATATGCAAACTTGTGTGC 109848
 |||||

RESULT 11

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

D0022571 550 bp DNA linear STS 28-MAY-2005
 Beta vulgaris chromosome 1, sequence tagged site.

D0022571.1 GI:66394762
 STS.

Beta vulgaris
 Beta vulgaris

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 Caryophyllales; Amaranthaceae; Beta.

1 (bases 1 to 550)
 Friesen,T.L., Weiland,J.J., Ashheim,M.L., Hunger,S., Borchardt,D.C.
 and Lewellen,R.T.

Identification of a SCAR marker associated with Bm the Beet mosaic
 virus resistance gene on chromosome 1

Unpublished
 2 (bases 1 to 550)
 Weiland,J.J. and Friesen,T.L.

Submitted (29-APR-2005) Sugarbeet and Potato Research,
 USDA-Agricultural Research Service, 1307 18th St. N, Fargo, ND
 58105, USA

FEATURES

Source

Location/Qualifiers

1..550
 /organism="Beta vulgaris"
 /mol_type="genomic DNA"
 /cultivar="C719"
 /db_xref="taxon:161934"
 /chromosome="1"
 /note="fwd primer seq: GAGTGAAGTGGTGAAGCAAGTGTG;
 rev primer seq: GGAAGCAATTAATCTTTATAGT;
 fwd_primer_name: Rbm05fwd; rev_primer_name: Rbm05rev"
 1..550
 /note="SCAR marker associated with resistance to Beet
 mosaic virus"

ORIGIN

Query Match 4.6%; Score 170.4; DB 10; Length 550;
 Best Local Similarity 75.0%; Pred. No. 2.7e-27;
 Matches 213; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 1361 TTGTATGATCTTGTGATGAGAGAGCAAGTATGCTTGAACGGATTTACGTCA 1420
 |||||

DB 96 TTGTATGATCTTGTGATGAGAGAGGTTCAAGTATGCTTGAACGGATTTACMAACA 155
 |||||

QY 1421 CAAATGATGAGCTTTGGAAGATCACTCCGGTTATGCTTGAAGCTACTTGGCTTA 1480
 |||||

DB 156 GAGGTGACGAGCAATTTGAGAGATGACACCGTTATGCTTGAAGCTTTGGCTTTG 215
 |||||

QY 1481 CCGCTGTGATGATTAACGCTGCGCAAAAAGACTAAATGTTTAAGCGGTGTGCGAATATT 1540
 |||||

DB 216 CTTCTCAATGATGATTTTTCGGAAGAAAAGATGAAGGCTTCAACGGGTGACGAATATA 275
 |||||

QY 1541 TTGTGCTGTGTTGAGAGGAGTGAAGCATCAGCTCTGTGTGGGGTTTGAACCGTGAAG 1600
 |||||

DB 276 TTATGCTGTGTTGAGAGGAGTGTGCGGACGTAATAGTGTGTTTACTCTGTGAGAT 335
 |||||

QY 1601 TTATGATGAGAGGCTTTTACGATGACAGCTGCTGAGCAGGT 1644
 |||||

DB 336 TTATGATGAGAGGCTTTTTCATATGACGACACTGACAGGT 379
 |||||

RESULT 12
 AC160013
 LOCUS 117818 bp DNA linear HTG 02-JUN-2005
 DEFINITION Medicago truncatula clone mtn2-162b23, WORKING DRAFT SEQUENCE, 4
 ordered pieces.
 ACCESSION AC160013
 VERSION AC160013.10 GI:66865008

VERSION AR579680.1 GI:56583130
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 1141)
Kunst, L. and Clemens, S.
Regulation of embryonic transcription in plants
Patent: US 6784342-A 22 31-AUG-2004;
The University of British Columbia; Vancouver;
CAN;
FEATURES
source location/Qualifiers
1..1141
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 1.5%; Score 54; DB 6; Length 1141;
Best Local Similarity 10.0%; Pred. No. 0.45;
Matches 79; Conservative 311; Mismatches 389; Indels 8; Gaps 2;
2393 TTAGGTGATGCTCTGTGTGTAACAATGATGAGCGCGTGGTGGAGCTGGTCTTATG 2452
67 KRRRYNNKSRWKKWYKKKWTBNCANNTSBRTHARWDMKTAYBMTNKGKTMGR 126
2453 CAGAAGCTGTAGAACCCTCTGAA-ACTTTGAACATAATGATTAATGCAATTCGAGCTGG 2511
127 YWRWABMDTYDHHVYTAAMNNAWTTMCMMDKDKRTTRWKKNNATGMDDDTKYHMMNN 186
2512 GTCTCAGAGATGACGCTTGATGAAACTACTGTTGAAATGTCCTGCTGATATGTTAAAG 2571
187 NGBCTVTMVRVYKTRDWSBKRMNYGMBWKMWSYDVLYWVMDMCKRKRVRWVTRG 246
2572 GAGGCAAGTGAAGATGCTGCTGTGTGCGATGCGAATGATGATTTGACTGTTTCAAGC 2631
247 RRRNYMAMBTHARRRNNNGMTBAMAAYRRMTNNNNNNNAKACRKAATGMRABVNSTC 306
2632 CAGAAGTATTTCTTAAAGCAGCTCATCTTTTCAAGCAAGATATGCTTTCTTATG 2691
307 TTWKSRTKVRTSCWANNCRAGDANKDHKMKMSAAGVYNNNNNNNNWTKKABHBAW 366
2692 GAATCTGATGCTGCTACCATAGATGATTAATGATGCAATTTCAATATCTGATG 2751
367 DMVHSAWKWNAHAHYSRKMTBYKRTIVNNNNNGTTMKRMWAMWYKMDMBGTYN 426
2752 CTCAAATATGCTGTTTGTGAGCTAAGAACTAGTCCCACTTAATACATGTCGCAA 2811
427 NNNNGRTTYGTTKKNKMTTYKKRANNCCKRAMDHKTCHNNTTWKMKTYMNNCYKS 486
2812 AGTTGTACCAAGATTAAACAAGTTGCTAGTAATTTCACTAATTATGCT-----GCTT 2864
487 MTNGKSHRBAALVYTWMMWRVYAHANNNNNDYMKACTYKKBVCCKMNNVAAWTK 546
2865 GAATTTTGTGATCAACCTGTAGACAGAAATGTAATTTCACTCTCAACATTTCTGTTAG 2924
547 SSMWTSRYRKYKTNNSWRMSDTRSMBGRANNYARABHYGKMTRWMBWMSHTWBHBAAG 606
2925 AATAACGTAGATGATGAGATGCTTAGTGCTGCTTTGCTCAACTTTCTTCTTGATT 2984
607 AAHYMBWMBYBAKCHCKAMYKAKKYTAGAGSGSNNNNNNNNNNNNNNNNATCARDYTAAS 666
2985 TTTTTCCTTTTCATTTAGGCTCAGTCAAGCTCAGATTCAGAACACTTCCCAAGATG 3044
667 RHYAANAANKWYVYKBAANNAVYTHANNMWGCMNNATDTRRTWKNNNNNAGTWKNNNN 726
3045 ATGCTAGACATGACAGAAATATGATTCAGAGTGCAGAAATGATTAAGTCTGCTGTTG 3104
727 NAKNSAAKNVYAAAVYKAKKRWANWYKMGWADABATTDKRNNGAYKYKTYTNNN 786
3105 GGCCTGATCAGCGCATAGAAATGTTACACAGAGTGAAGGGAATTAATTCACATTCATCA 3164
787 NTYRGVVTNTAARDWANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNAY 846

QY 3165 ATTGTGT 3171
DB 847 AAWTNKW 853
RESULT 15
LOCUS AX083744 1141 bp DNA linear PAT 28-FEB-2001
DEFINITION Sequence 22 from Patent WO0111061.
ACCESSION AX083744
VERSION AX083744.1 GI:13185472
KEYWORDS
SOURCE
ORGANISM
1
Kunst, L. and Clemens, S.
Regulation of embryonic transcription in plants
Patent: WO 0111061-A 22 15-FEB-2001;
UNIVERSITY OF BRITISH COLUMBIA (CA)
FEATURES
source location/Qualifiers
1..1141
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
1..1141
/note="consensus sequence of A.c., L.a., and B.n. FAE1 promoters"
ORIGIN
Query Match 1.5%; Score 54; DB 6; Length 1141;
Best Local Similarity 10.0%; Pred. No. 0.45;
Matches 79; Conservative 311; Mismatches 389; Indels 8; Gaps 2;
2393 TTAGGTGATGCTCTGTGTGTAACAATGATGAGCGCGTGGTGGAGCTGGTCTTATG 2452
67 KRRRYNNKSRWKKWYKKKWTBNCANNTSBRTHARWDMKTAYBMTNKGKTMGR 126
2453 CAGAAGCTGTAGAACCCTCTGAA-ACTTTGAACATAATGATTAATGCAATTCGAGCTGG 2511
127 YWRWABMDTYDHHVYTAAMNNAWTTMCMMDKDKRTTRWKKNNATGMDDDTKYHMMNN 186
2512 GTCTCAGAGATGACGCTTGATGAAACTACTGTTGAAATGTCCTGCTGATATGTTAAAG 2571
187 NGBCTVTMVRVYKTRDWSBKRMNYGMBWKMWSYDVLYWVMDMCKRKRVRWVTRG 246
2572 GAGGCAAGTGAAGATGCTGCTGTGTGCGATGCGAATGATGATTTGACTGTTTCAAGC 2631
247 RRRNYMAMBTHARRRNNNGMTBAMAAYRRMTNNNNNNNAKACRKAATGMRABVNSTC 306
2632 CAGAAGTATTTCTTAAAGCAGCTCATCTTTTCAAGCAAGATATGCTTTCTTATG 2691
307 TTWKSRTKVRTSCWANNCRAGDANKDHKMKMSAAGVYNNNNNNNNWTKKABHBAW 366
2692 GAATCTGATGCTGCTACCATAGATGATTAATGATGCAATTTCAATATCTGATG 2751
367 DMVHSAWKWNAHAHYSRKMTBYKRTIVNNNNNGTTMKRMWAMWYKMDMBGTYN 426
2752 CTCAAATATGCTGTTTGTGAGCTAAGAACTAGTCCCACTTAATACATGTCGCAA 2811
427 NNNNGRTTYGTTKKNKMTTYKKRANNCCKRAMDHKTCHNNTTWKMKTYMNNCYKS 486
2812 AGTTGTACCAAGATTAAACAAGTTGCTAGTAATTTCACTAATTATGCT-----GCTT 2864
487 MTNGKSHRBAALVYTWMMWRVYAHANNNNNDYMKACTYKKBVCCKMNNVAAWTK 546
2865 GAATTTTGTGATCAACCTGTAGACAGAAATGTAATTTCACTCTCAACATTTCTGTTAG 2924
547 SSMWTSRYRKYKTNNSWRMSDTRSMBGRANNYARABHYGKMTRWMBWMSHTWBHBAAG 606
2925 AATAACGTAGATGATGAGATGCTTAGTGCTGCTTTGCTCAACTTTCTTCTTGATT 2984
607 AAHYMBWMBYBAKCHCKAMYKAKKYTAGAGSGSNNNNNNNNNNNNNNNNNNNNNNNNNN 666

OY 2985 TTTTCTTTTCGATTAGGTCAGTCAAGCTGACGATTCAGAGCACTCCAGATG 3044
DB 667 RMYAANAKMYYYKBAANNAYYTHANNWGCWNNATDTRRTWKNNNNNAGTWKNNNN 726
OY 3045 ATGCTAGACTGCAGAGATATAGTATCCAGTGCAGAGATTAACTCTGCTTTG 3104
DB 727 NAKNLSAAKNYAAAVKAKKMRWANKWAMRGWHDAAABTTDKRNNGAYTKYTTTNN 786
OY 3105 GGCCTGATCACCGCATAGAAATGTTACCGAGGTGAGGGAATTAATCTACAATCATCA 3164
DB 787 NTYRGVTTTARDDGMAANNNNNNNNNNNGWSDWVTWVAANYGTNNNNNNNNNAY 846
OY 3165 ATYGTGT 3171
DB 847 AMWTNKW 853

Search completed: December 11, 2005, 03:43:31
Job time : 18184.9 secs


```

; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C128
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 60/059111
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 509
; LENGTH: 1281
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-131-826A-509

Query Match      1.8%; Score 43; DB 6; Length 1281;
Best Local Similarity 47.6%; Pred. No. 0.0033;
Matches 127; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

QY 421 TCTATCTCGGTCTAGAGAGAGATGACATGAGGCTCTTGTGATGATGAAGAAGCTTCA 480
DB 581 TTTGATCTCTTCTTGTGCTTCAAGATGATGTTGTGATGTTGATGATGAAGACAGCA 522
QY 481 GTCATCACTGATGTTCTTGGATTAAGGTTCTTGGGGCTCTCTGTATGTCAGAAAGCT 540
DB 521 GATGTCATGACATGATTTGGGTGTAAGCTTGAAGGTGATGATGATGATGATGATGAT 462
QY 541 GGTGAGACTGAGATAGTTCTTCTGGGTGAGGCTCTTCTTGAAGAGAGGTTGCTTAAG 600
DB 461 TGTGAACACTGTTGTTTGGGTGTAAGCTTGAAGGTGATGATGATGATGATGATGATG 402
QY 601 TCGTTAAGCAAGATGTTGTTTATGTTATGAGCGCTTGGCTTCTGATGATCTCGAGGAT 660
DB 401 ACCATCCCTGTTGTTGTTATTAATGATCCGCTGATGATGATGATGATGATGATGATG 342
QY 661 GCTATGACATGATGATCAACCTGATTTT 687
DB 341 TTACTGAGACTGAGGCACTGAACTT 315

RESULT 3
US-11-121-086-61/c
; Sequence 61, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
```

```

; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121.086
; PRIOR FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 61
; LENGTH: 169495
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (70072)..(70171)
; OTHER INFORMATION: a, c, g, t, unknown or other
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (139457)..(157244)
; OTHER INFORMATION: a, c, g, t, unknown or other
US-11-121-086-61

Query Match      1.8%; Score 42.4; DB 7; Length 169495;
Best Local Similarity 48.7%; Pred. No. 0.44;
Matches 115; Conservative 0; Mismatches 121; Indels 0; Gaps 0;

QY 1657 GTCTTAGTGTAGTCTCTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1716
DB 67566 GTGTTGCTGCTGATGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 67507
QY 1717 ATTGCAAGCTGTGAAGACCTCTGAAACTTGAACATGATTTATGCAATTCAGACT 1776
DB 67506 ACTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 67447
QY 1777 GGGGTCTCAGAGTACGTTGATGAAACTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1836
DB 67446 GCTGTTGATGATGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 67387
QY 1837 AAGAGGCAAGTGTGAAGATCTTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1892
DB 67386 ATTGCTCTTCTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 67331

RESULT 4
US-10-829-826B-21
; Sequence 21, Application US/10829826B
; Publication No. US20050266397A1
; GENERAL INFORMATION:
; APPLICANT: Ecker, David J.
; APPLICANT: Hotstadler, Steven A.
; APPLICANT: Sampath, Rangarajan
; APPLICANT: Blyn, Lawrence B.
; APPLICANT: Hall, Thomas A.
; APPLICANT: Massire, Christian
; TITLE OF INVENTION: METHOD FOR IDENTIFICATION OF CORONAVIRUSES
; FILE REFERENCE: IBIS0075-100 (DIBIS-0058US)
; CURRENT APPLICATION NUMBER: US/10/829.826B
; CURRENT FILING DATE: 2004-04-22
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 31028
; TYPE: DNA
; ORGANISM: Coronavirus
US-10-829-826B-21

Query Match      1.7%; Score 40.2; DB 6; Length 31028;
Best Local Similarity 51.4%; Pred. No. 0.63;
Matches 99; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 1651 GAGACAGTGTATTAGTGTAGATCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1710
```

```
Db 3700 GATGCTAATGCTTTTAAAGGATGATGTCACATGTCAGAGTGTGATGCTAATG 3759
Qy 1711 GTCTTAAATGAGAAAGCTGTAAGACCCCTGAAACTTGAACATAATGATGCAAT 1770
Db 3760 GTACTTAAATGATGATGATGTCATTAACAGCCCACTTGTCTTAAGATAAGTTGTT 3819
Qy 1771 CGAGCTGGGGTCTCAGAGAGTAGACGTTGATGAACTACTGTGAAATGTCCGTTGCTGAT 1830
Db 3820 TGTGCAATTAATTAAGAGGCTAGTGTGATTAAGACGCTGTGTGTGATGATTAATGAT 3879
Qy 1831 A 1831
Db 3880 A 3880
```

RESULT 5
US-10-829-826B-22

```
; Sequence 22, Application US/10829826B
; Publication No. US20050266397A1
; GENERAL INFORMATION:
; APPLICANT: Becker, David J.
; APPLICANT: Hofstadler, Steven A.
; APPLICANT: Sampath, Rangarajan
; APPLICANT: Blyn, Lawrence B.
; APPLICANT: Hall, Thomas A.
; APPLICANT: Massire, Christian
; TITLE OF INVENTION: METHOD FOR IDENTIFICATION OF CORONAVIRUSES
; FILE REFERENCE: IBIS0075-100 (DIBIS-0058US)
; CURRENT APPLICATION NUMBER: US/10/829,826B
; CURRENT FILING DATE: 2004-04-22
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 31028
; TYPE: DNA
; ORGANISM: Coronavirus
US-10-829-826B-22
```

Query Match 1.7%; Score 40.2; DB 6; Length 31028;

Best Local Similarity 51.4%; Pred. No. 0.63; Matches 93; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

```
Qy 1651 GAGACAGTGTAGTATGATCCTGTGTAACAATGTAAGCCGTGATGAGCCCTGAT 1710
Db 3700 GATGCTAATGCTTTTAAAGGATGATGTCACATGTCAGAGTGTGATGCTAATG 3759
Qy 1711 GTCTTAAATGAGAAAGCTGTAAGACCCCTGAAACTTGAACATAATGATGCAAT 1770
Db 3760 GTACTTAAATGATGATGATGTCATTAACAGCCCACTTGTCTTAAGATAAGTTGTT 3819
Qy 1771 CGAGCTGGGGTCTCAGAGAGTAGACGTTGATGAACTACTGTGAAATGTCCGTTGCTGAT 1830
Db 3820 TGTGCAATTAATTAAGAGGCTAGTGTGATTAAGACGCTGTGTGTGATGATTAATGAT 3879
Qy 1831 A 1831
Db 3880 A 3880
```

RESULT 6
US-10-829-826B-26

```
; Sequence 26, Application US/10829826B
; Publication No. US20050266397A1
; GENERAL INFORMATION:
; APPLICANT: Becker, David J.
; APPLICANT: Hofstadler, Steven A.
; APPLICANT: Sampath, Rangarajan
; APPLICANT: Blyn, Lawrence B.
; APPLICANT: Hall, Thomas A.
; APPLICANT: Massire, Christian
; TITLE OF INVENTION: METHOD FOR IDENTIFICATION OF CORONAVIRUSES
; FILE REFERENCE: IBIS0075-100 (DIBIS-0058US)
```

```
; CURRENT APPLICATION NUMBER: US/10/829,826B
; CURRENT FILING DATE: 2004-04-22
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 31028
; TYPE: DNA
; ORGANISM: Coronavirus
US-10-829-826B-26
```

Query Match 1.7%; Score 40.2; DB 6; Length 31028;

Best Local Similarity 51.4%; Pred. No. 0.63; Matches 93; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

```
Qy 1651 GAGACAGTGTAGTATGATCCTGTGTAACAATGTAAGCCCGTATGAGCCCTGAT 1710
Db 3700 GATGCTAATGCTTTTAAAGGATGATGTCACATGTCAGAGTGTGATGCTAATG 3759
Qy 1711 GTCTTAAATGAGAAAGCTGTAAGACCCCTGAAACTTGAACATAATGATGCAAT 1770
Db 3760 GTACTTAAATGATGATGATGTCATTAACAGCCCACTTGTCTTAAGATAAGTTGTT 3819
Qy 1771 CGAGCTGGGGTCTCAGAGAGTAGACGTTGATGAACTACTGTGAAATGTCCGTTGCTGAT 1830
Db 3820 TGTGCAATTAATTAAGAGGCTAGTGTGATTAAGACGCTGTGTGTGATGATTAATGAT 3879
Qy 1831 A 1831
Db 3880 A 3880
```

RESULT 7
US-10-829-826B-24

```
; Sequence 24, Application US/10829826B
; Publication No. US20050266397A1
; GENERAL INFORMATION:
; APPLICANT: Becker, David J.
; APPLICANT: Hofstadler, Steven A.
; APPLICANT: Sampath, Rangarajan
; APPLICANT: Blyn, Lawrence B.
; APPLICANT: Hall, Thomas A.
; APPLICANT: Massire, Christian
; TITLE OF INVENTION: METHOD FOR IDENTIFICATION OF CORONAVIRUSES
; FILE REFERENCE: IBIS0075-100 (DIBIS-0058US)
; CURRENT APPLICATION NUMBER: US/10/829,826B
; CURRENT FILING DATE: 2004-04-22
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 31100
; TYPE: DNA
; ORGANISM: Coronavirus
US-10-829-826B-24
```

Query Match 1.7%; Score 40.2; DB 6; Length 31100;

Best Local Similarity 51.4%; Pred. No. 0.63; Matches 93; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

```
Qy 1651 GAGACAGTGTAGTATGATCCTGTGTAACAATGTAAGCCCGTATGAGCCCTGAT 1710
Db 3700 GATGCTAATGCTTTTAAAGGATGATGTCACATGTCAGAGTGTGATGCTAATG 3759
Qy 1711 GTCTTAAATGAGAAAGCTGTAAGACCCCTGAAACTTGAACATAATGATGCAAT 1770
Db 3760 GTACTTAAATGATGATGATGTCATTAACAGCCCACTTGTCTTAAGATAAGTTGTT 3819
Qy 1771 CGAGCTGGGGTCTCAGAGAGTAGACGTTGATGAACTACTGTGAAATGTCCGTTGCTGAT 1830
Db 3820 TGTGCAATTAATTAAGAGGCTAGTGTGATTAAGACGCTGTGTGTGATGATTAATGAT 3879
Qy 1831 A 1831
Db 3880 A 3880
```



```
Query Match 1.6%; Score 39.2; DB 7; Length 1317;
Best Local Similarity 51.7%; Pred. No. 0.14;
Matches 89; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 34 TCCCATTCATTTATGCCGATTACACCGGCGACGACCAAAAGTCCGAGTACCAAC 93
DB 129 TCCACACCCACGCGTTGAAGAACAGCCACCGCCGCGCCACCTCCACCTCCAC 188
QY 94 ACCTTACAACTATCTGCTCCGCGACGAATGGCGGACCGCTCTCTCCGACTTCAT 153
DB 189 ACCACTCTCTCTCTCTCTCTCCACCGAGTACAGCCAGCAAGTTGAAGAGGCTCCCC 248
QY 154 TTCACTCCGATTCCTCTCTCTCTCTCTCTGCGACCGCGACCAACCGCCCA 205
DB 249 TCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300

RESULT 12
US-11-140-417-22
; Sequence 22, Application US/11140417
; Publication No. US20050266515A1
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Deborah A
; APPLICANT: Edgy, Edward M
; TITLE OF INVENTION: GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE-S (GAPDS), A GLYCOLYTIC
; TITLE OF INVENTION: ENZYME EXPRESSED ONLY IN MALE GERM CELLS, IS A TARGET FOR MLE
; TITLE OF INVENTION: CONTRACEPTION
; FILE REFERENCE: 421/76/2 PCT/CIP
; CURRENT APPLICATION NUMBER: US/11/140,417
; CURRENT FILING DATE: 2005-05-27
; PRIOR APPLICATION NUMBER: US 60/429,638
; PRIOR FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: PCT/US2003/037800
; PRIOR FILING DATE: 2003-11-26
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 22
; LENGTH: 11462
; TYPE: DNA
; ORGANISM: Mus musculus
; US-11-140-417-22

Query Match 1.6%; Score 39.2; DB 7; Length 11462;
Best Local Similarity 51.7%; Pred. No. 0.64;
Matches 89; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 34 TCCCATTCATTTATGCCGATTACACCGGCGACGACCAAAAGTCCGAGTACCAAC 93
DB 3221 TCCACACCCACGCGTTGAAGAACAGCCACCGCCGCGCCACCTCCACCTCCAC 3280
QY 94 ACCTTACAACTATCTGCTCCGCGACGAATGGCGGACCGCTCTCTCCGACTTCAT 153
DB 3281 ACCACTCTCTCTCTCTCTCTCCACCGAGTACAGCCAGCAAGTTGAAGAGGCTCCCC 3340
QY 154 TTCACTCCGATTCCTCTCTCTCTCTCTCTGCGACCGCGACCAACCGCCCA 205
DB 3341 TCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3392

RESULT 13
US-11-121-086-24
; Sequence 24, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138, 6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
```

```
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 24
; LENGTH: 120096
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-11-121-086-24

Query Match 1.6%; Score 39.2; DB 7; Length 120096;
Best Local Similarity 54.1%; Pred. No. 3.3;
Matches 80; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 93 CACCTTACAACTATCTGCTCCGCGACGAATGGCGGACCGCTCTCTCCGACTTC 152
DB 41777 CTCCTCCATTTATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 41836
QY 153 TTTCACCTCCGATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 212
DB 41837 CTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 41896
QY 213 CTCCTGCGACATCTATTGATCGTCC 240
DB 41897 TTCTTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 41924

RESULT 14
US-11-121-086-5/c
; Sequence 5, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138, 6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 5
; LENGTH: 153376
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-11-121-086-5

Query Match 1.6%; Score 39; DB 7; Length 153376;
Best Local Similarity 51.4%; Pred. No. 4.5;
Matches 90; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

QY 32 TCTCCCATTCATTTATGCCGATTACACCGGCGACGACCAAAAGTCCGAGTACCA 91
DB 15887 TCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 15828
QY 92 ACACCTTACAACTATCTGCTCCGCGACGAATGGCGGACCGCTCTCTCTCCGACTTC 151
DB 15827 GCACCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 15768
QY 152 ATTTCACCTCCGATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 206
DB 15767 CCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 15713

RESULT 15
US-11-167-856-23
; Sequence 23, Application US/11167856
; Publication No. US20050268352A1
; GENERAL INFORMATION:
; APPLICANT: Nikolaev, Basil J
; APPLICANT: Murtele, Eve S
; APPLICANT: Oliver, David J
; APPLICANT: Behal, Robert
; APPLICANT: Schnable, Patrick S
```

APPLICANT: Ke, Jinsan
APPLICANT: Johnson, Jerry L
APPLICANT: Allred, Carolyn C
APPLICANT: Fatiand, Beth
APPLICANT: Lutziger, Isabelle
APPLICANT: Wen, Tsui-Jung
TITLE OF INVENTION: Materials and Methods for the Alteration of Enzyme and
FILE REFERENCE: P2194USDIV-2
CURRENT APPLICATION NUMBER: US/11/167,856
CURRENT FILING DATE: 2005-06-27
PRIOR APPLICATION NUMBER: US 10/293,865
PRIOR FILING DATE: 2002-11-13
PRIOR APPLICATION NUMBER: US 09/344,882
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: US 60/090,717
PRIOR FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 38
SOFTWARE: PatentIn Ver. 3.1
SEQ ID NO 23
LENGTH: 1772
TYPE: DNA
ORGANISM: Arabidopsis Thaliana
FEATURE:
NAME/KEY: exon
LOCATION: (1)..(834)
FEATURE:
NAME/KEY: exon
LOCATION: (836)..(930)
FEATURE:
NAME/KEY: exon
LOCATION: (932)..(939)
FEATURE:
NAME/KEY: exon
LOCATION: (941)..(1772)
US-11-167-856-23

Query Match 1.6%; Score 38.8; DB 7; Length 1772;
Best Local Similarity 46.6%; Pred. No. 0.23;
Matches 124; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

QY	832	CCGCTTGATGATTAACCGCGAAAGACTAATGTTTAAAGCGGTGTGCGAATATT	891
DB	705	CCTATCTGCTCTTCTGTGGAACTACTTCAATGAGGCTGACTTCTGATGAGATT	764
QY	892	TTGTGCTGTTGAGAGAGTGAAGCATCAGCTTTGTTGGGGTTTGACCCGTGAAG	951
DB	765	GTGAATATAGTTTCTGGAATTTGGGCTACTGCTGTGCACTATAGTACATGAGC	824
QY	952	TTTATGAATGAGCGCTTTTACGAATGACAGCTGCTGAGCAGGTTGATCTTTTGTAGCT	1011
DB	825	GTTGATTAAGGTTGCTTTCACCGGCTACTGATGTTGGGAAGATTATCTTGAAGTTAGCT	884
QY	1012	ACCCAGCAATATCCAGCAGATGATTTGAAGTTTACGAAGTTGCACTGCTTTGTG	1071
DB	885	TCAAAAAGCAACCTTAAGCAGTACTCTTGAGCTGAGGAAGTCACCAATTCATTGTGA	944
QY	1072	GCTCAGCTTTTATTTGTTAAGGCC	1097
DB	945	TGTGAAGATGCTGATGTGATCAGGC	970

Search completed: December 10, 2005, 18:24:03
Job time : 208.429 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: December 10, 2005, 06:08:12 ; Search time 1835.5 Seconds
(without alignments)
10839.614 Million cell updates/sec

Title: US-10-600-070b-1

Perfect score: 2406
Sequence: 1 atggaagctctgagtcacgt.....gctctgtcttcgcatcataa 2406

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA_Main:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2406	100.0	2406	US-10-600-070-1	Sequence 1, Appli
2	2402.8	99.9	2406	US-10-600-070-9	Sequence 9, Appli
3	2399.6	99.7	2406	US-10-600-070-128	Sequence 128, App
4	2399.6	99.7	2637	US-10-600-070-130	Sequence 130, App
5	2399.6	99.7	2637	US-10-739-930-227	Sequence 227, App
6	1724	71.7	3667	US-10-600-070-3	Sequence 3, Appli
7	1720.8	71.5	3667	US-10-600-070-10	Sequence 10, Appli
8	482	20.0	561	US-10-600-070-132	Sequence 132, App
9	481	20.0	2283	US-10-600-070-126	Sequence 126, App
10	472.8	19.7	1146	US-10-424-559-129007	Sequence 129007,
11	361.4	15.0	1411	US-10-425-115-81853	Sequence 81853, A
12	328.8	13.7	631	US-10-600-070-184	Sequence 184, App
13	278.8	11.6	660	US-10-600-070-135	Sequence 135, App
14	231.6	9.6	1039	US-10-424-559-35059	Sequence 35059, A
15	224	9.3	537	US-10-600-070-143	Sequence 143, App
16	223.4	9.2	552	US-10-021-323-3536	Sequence 3636, App
17	207	8.6	545	US-10-600-070-175	Sequence 175, App
18	197.2	8.2	491	US-10-600-070-114	Sequence 174, App
19	195.8	8.1	467	US-10-600-070-185	Sequence 185, App
20	193.8	8.1	460	US-10-600-070-187	Sequence 187, App
21	190	7.9	446	US-10-600-070-151	Sequence 151, App
22	188.8	7.8	1703	US-10-437-963-69933	Sequence 69933, A
23	186.6	7.8	652	US-10-600-070-186	Sequence 186, App

24	178.6	7.4	607	7	US-10-021-323-3562	Sequence 3562, App
25	175	7.3	439	3	US-09-732-627A-2154	Sequence 2154, App
26	170.4	7.1	608	7	US-10-600-070-137	Sequence 137, App
27	170.2	7.1	563	7	US-10-600-070-178	Sequence 178, App
28	169.8	7.1	622	7	US-10-600-070-146	Sequence 146, App
29	169	7.0	1536	8	US-10-425-115-57452	Sequence 57452, A
30	167.6	6.9	1032	7	US-10-767-701-9474	Sequence 9474, App
31	165.4	6.9	527	7	US-10-600-070-134	Sequence 134, App
32	162	6.7	647	7	US-10-767-701-4069	Sequence 4069, App
33	158.6	6.6	871	7	US-10-600-070-153	Sequence 153, App
34	155.4	6.5	307	7	US-10-600-070-138	Sequence 138, App
35	150.4	6.3	418	7	US-10-600-070-144	Sequence 144, App
36	146.2	6.1	2130	7	US-10-437-963-69932	Sequence 69932, A
37	144.6	6.0	653	7	US-10-600-070-148	Sequence 148, App
38	144.2	6.0	479	7	US-10-600-070-150	Sequence 150, App
39	130.8	5.4	420	7	US-10-600-070-176	Sequence 176, App
40	128.6	5.3	480	7	US-10-600-070-145	Sequence 145, App
41	127.6	5.3	535	7	US-10-600-070-149	Sequence 149, App
42	124.2	5.2	360	7	US-10-600-070-179	Sequence 179, App
43	118.8	4.9	336	7	US-10-600-070-142	Sequence 142, App
44	118.6	4.9	541	7	US-10-600-070-154	Sequence 154, App
45	117.8	4.9	309	7	US-10-600-070-141	Sequence 141, App

ALIGNMENTS

RESULT 1
US-10-600-070-1
; Sequence 1, Application US/10600070
; Publication No. US20040139500A1
; GENERAL INFORMATION:
; APPLICANT: Oeteryoung, Katherine W.
; APPLICANT: Vitha, Stanislav
; APPLICANT: Kokscharova, Olga A.
; APPLICANT: Gao, Honggo
; TITLE OF INVENTION: Placid Division and Related Genes and Proteins, and Methods of
; FILE REFERENCE: MSU-08153
; CURRENT APPLICATION NUMBER: US/10/600,070
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 2406
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-600-070-1

Query Match	100.0%;	Score 2406;	DB 7;	Length 2406;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2406;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ATGGAAGCTCTGAGTCAAGTCGATGCTCTCTCCCATTCGAATTAATGCGATTACCA	60	
DB	1	ATGGAAGCTCTGAGTCAAGTCGATGCTCTCTCCCATTCGAATTAATGCGATTACCA	60	
QY	61	CCGGCGAAGCAAGAGCTCCGAGTGAAGCAAGCTCTTAACATATCTGTCGGCAGC	120	
DB	61	CCGGCGAAGCAAGAGCTCCGAGTGAAGCAAGCTCTTAACATATCTGTCGGCAGC	120	
QY	121	AAATGGGCGGACGCTCTCTCTCCGACTTCAATTTCACTCCGATTCCTCTCTCTCC	180	
DB	121	AAATGGGCGGACGCTCTCTCTCCGACTTCAATTTCACTCCGATTCCTCTCTCTCC	180	
QY	181	TTCCGCAACCGGACACCAACGCGCATCTGCTCTCTGCGACCATTAATGATGATGCC	240	
DB	181	TTCCGCAACCGGACACCAACGCGCATCTGCTCTCTGCGACCATTAATGATGATGCC	240	
QY	241	GAAAGCCAGGCGCCATCCCATTAATTTACAGGATTAATGAGGCTCAACAACATTTTC	300	
DB	241	GAAAGCCAGGCGCCATCCCATTAATTTACAGGATTAATGAGGCTCAACAACATTTTC	300	

Qy	301	TTAACGATGGAATCAGAAAGACATTGAAAGCTAGGCTTTGAAACCGCCGCAATTGGT	360	Db	1381	TTTTTTTAACTCGGGGACTACTATGATGATCCATATGCTTTGAGTTACTTGAAAGAGTG	1440
Db	301	TTAACCGATGGAATCAGAAAGACATTGAAAGCTAGGCTTTGAAACCGCCGCAATTGGT	360	Qy	1441	AAATTTAACTCGGGGACTACTATGATGATCCATATGCTTTGAGTTACTTGAAAGAGTG	1500
Qy	361	TTACGCGAGACGCTTTAATCAGCCGAGACAGATTCTTCAAGCTGCTTGGAACCTTG	420	Db	1441	AAATTTAACTCGGGGACTACTATGATGATCCATATGCTTTGAGTTACTTGAAAGAGTG	1500
Db	361	TTACGCGAGACGCTTTAATCAGCCGAGACAGATTCTTCAAGCTGCTTGGAACCTTG	420	Qy	1501	GAGTAGTTCAAGGTTCTCTTTAGCTGCTGCACTATGCGAAGATTTGAGCCGAG	1560
Qy	421	TCTAATCTCGGTCTAGAAAGAGATACAAATGAGTCTTCTGATGATGAAAGACTACA	480	Db	1501	GAGTAGTTCAAGGTTCTCTTTAGCTGCTGCACTATGCGAAGATTTGAGCCGAG	1560
Db	421	TCTAATCTCGGTCTAGAAAGAGATACAAATGAGTCTTCTGATGATGAAAGACTACA	480	Qy	1561	CATGTGAAGCTAGTGTCTATGCAAGGCACTGCAAGAGTTTCTTCCCTGATACGAT	1620
Qy	481	GTCATCACTGATGTTCTTGAGATAAGGTTCTGAGGCTCTCTGTGTAATGCAAGAGT	540	Db	1561	CATGTGAAGCTAGTGTCTATGCAAGGCACTGCAAGAGTTTCTTCCCTGATACGAT	1620
Db	481	GTCATCACTGATGTTCTTGAGATAAGGTTCTGAGGCTCTCTGTGTAATGCAAGAGT	540	Qy	1621	AGAACTCGGCTGAAACCAAGAGATGTCAGAGACAGTGTATGATAGATCCGTTGGT	1680
Qy	541	GGTAGACTGAGATAGTCTTGAGGATTAAGGTTCTGAGGCTCTCTGTGTAATGCAAGAGT	600	Db	1621	AGAACTCGGCTGAAACCAAGAGATGTCAGAGACAGTGTATGATAGATCCGTTGGT	1680
Db	541	GGTAGACTGAGATAGTCTTGAGGATTAAGGTTCTGAGGCTCTCTGTGTAATGCAAGAGT	600	Qy	1681	AACAATGTAAGCCGTAATGATGAGCTGATGCTTTATGTCAGAAAGCTGTAAAGCCCTCT	1740
Qy	601	TCGTTTAAGCAAGAATGTTAGTTATGAGCGCTTGCGTTCTGATGTCGAGGAT	660	Db	1681	AACAATGTAAGCCGTAATGATGAGCTGATGCTTTATGTCAGAAAGCTGTAAAGCCCTCT	1740
Db	601	TCGTTTAAGCAAGAATGTTAGTTATGAGCGCTTGCGTTCTGATGTCGAGGAT	660	Qy	1741	GAATACTTTGAACCTAATGATTAATGCAATTCGAGCTGGGCTTCAGAGATAGCGTTGAT	1800
Qy	661	GCTATGCAATTGATCCACTGATTTTAATCTGTTATGAGTTGTTGAGAAAGCTTTG	720	Db	1741	GAATACTTTGAACCTAATGATTAATGCAATTCGAGCTGGGCTTCAGAGATAGCGTTGAT	1800
Db	661	GCTATGCAATTGATCCACTGATTTTAATCTGTTATGAGTTGTTGAGAAAGCTTTG	720	Qy	1801	GAATCTACTGTTGAAATGTCGTTGCTGATATGTTAAAGAGGCAAGTGAAGATCTTA	1860
Qy	721	AAGCTTTTACAGAGAGAGCAAGTAGCCCTTGACCGGATTTACGTCACAAATGAT	780	Db	1801	GAATCTACTGTTGAAATGTCGTTGCTGATATGTTAAAGAGGCAAGTGAAGATCTTA	1860
Db	721	AAGCTTTTACAGAGAGAGCAAGTAGCCCTTGACCGGATTTACGTCACAAATGAT	780	Qy	1861	GCTGCTGCTGTCGCAATTTGATGATTTCACTGTTCAGCCAGAGATATTTTCTTAAAGC	1920
Qy	781	GAGACTTTGGAAGAGATCACTCCGCTTATGCTTTGAGACTACTTGCGCTTACCGCTGGT	840	Db	1861	GCTGCTGCTGTCGCAATTTGATGATTTCACTGTTCAGCCAGAGATATTTTCTTAAAGC	1920
Db	781	GAGACTTTGGAAGAGATCACTCCGCTTATGCTTTGAGACTACTTGCGCTTACCGCTGGT	840	Qy	1921	AGCTCATCTTTTCAACCCAGAGATATGCTTTCTTAATGAAATCTGATGTCGTAACATA	1980
Qy	841	GATGATTAACGCTGCGAAAAGACTAAATGCTTTAAGCGGTGTCGGAATATTTTGCTCT	900	Db	1921	AGCTCATCTTTTCAACCCAGAGATATGCTTTCTTAATGAAATCTGATGTCGTAACATA	1980
Db	841	GATGATTAACGCTGCGAAAAGACTAAATGCTTTAAGCGGTGTCGGAATATTTTGCTCT	900	Qy	1981	GGGTCACTGAGAGCTGACGATTCAGAAAGCACTTCCGAAATGATGCTGAAGCTGACAG	2040
Qy	901	GTTGAGAGAGGTGAGCATCAAGCTCTTGTTGGGGTTTGACCCGTAAGATTTATGAT	960	Db	1981	GGGTCACTGAGAGCTGACGATTCAGAAAGCACTTCCGAAATGATGCTGAAGCTGACAG	2040
Db	901	GTTGAGAGAGGTGAGCATCAAGCTCTTGTTGGGGTTTGACCCGTAAGATTTATGAT	960	Qy	2041	AAATATGATTCAAAGTGGCAAGAGATTAAGTCTTGCTTTTGCGCTGATCAACCGCAT	2100
Qy	961	GAGGCGTTTTACGATGACAGCTGCTGAGCAGGTTGATCTTTTGTAGTACCCCAAGC	1020	Db	2041	AAATATGATTCAAAGTGGCAAGAGATTAAGTCTTGCTTTTGCGCTGATCAACCGCAT	2100
Db	961	GAGGCGTTTTACGATGACAGCTGCTGAGCAGGTTGATCTTTTGTAGTACCCCAAGC	1020	Qy	2101	GAATGTTTACCAAGGTTTTGATGAGCGCAATGCTGAAGATTTGGAAGTTCGACAGCAGCT	2160
Qy	1021	AATATTCAGACAGAGTCAATTGAAATTAAGATTTGCAAGTTGCACTTGCTTGTCGCTCAAGCT	1080	Db	2101	GAATGTTTACCAAGGTTTTGATGAGCGCAATGCTGAAGATTTGGAAGTTCGACAGCAGCT	2160
Db	1021	AATATTCAGACAGAGTCAATTGAAATTAAGATTTGCAAGTTGCACTTGCTTGTCGCTCAAGCT	1080	Qy	2161	GAATCTGCGAGCTGGGTTGCTTTATGATTAATACATGTTGAAACATATCTGTTGACAGT	2220
Qy	1081	TTTATTTGTAAGAAGCCACCTTTTACAGAGTGTGATTAAGCAATTCAGCAACTTGAG	1140	Db	2161	GAATCTGCGAGCTGGGTTGCTTTATGATTAATACATGTTGAAACATATCTGTTGACAGT	2220
Db	1081	TTTATTTGTAAGAAGCCACCTTTTACAGAGTGTGATTAAGCAATTCAGCAACTTGAG	1140	Qy	2221	GTCACAGTCTCAGAGATGGAACCCGTCGCTGTGGAACAACCTGAGAGAGTCTGCT	2280
Qy	1141	CAGGCTAAGGTAAATGCTAATGAGATTCTGCGATGTTGTAATGATACACGGAATTAATGG	1200	Db	2221	GTCACAGTCTCAGAGATGGAACCCGTCGCTGTGGAACAACCTGAGAGAGTCTGCT	2280
Db	1141	CAGGCTAAGGTAAATGCTAATGAGATTCTGCGATGTTGTAATGATACACGGAATTAATGG	1200	Qy	2281	TGCTATCTGATTTGGTTCAATCCAGAAACAAATGCTATGATGTGCAAGACCTACACAACA	2340
Qy	1201	GAGATAGACTTCGCTTGAAAGAGGGAATCTGTGCACTGCTTAATGCAAGATTTGATGAA	1260	Db	2281	TGCTATCTGATTTGGTTCAATCCAGAAACAAATGCTATGATGTGCAAGACCTACACAACA	2340
Db	1201	GAGATAGACTTCGCTTGAAAGAGGGAATCTGTGCACTGCTTAATGCAAGATTTGATGAA	1260	Qy	2341	AGATACGAAGTTTCTGCTCAAGTCAAGGCTGGAATATCTGAAGGCTCTGTTCTTGCA	2400
Qy	1261	TGCGGTATGATGAGGCTTGAAGAGATTCACAATATAGGAATTCAGCTATTTGTG	1320	Db	2341	AGATACGAAGTTTCTGCTCAAGTCAAGGCTGGAATATCTGAAGGCTCTGTTCTTGCA	2400
Db	1261	TGCGGTATGATGAGGCTTGAAGAGATTCACAATATAGGAATTCAGCTATTTGTG	1320	Qy	2401	TCATTA 2406	
Qy	1321	GAGTTTGTGTTGAGAAATTCGAATCTGATGATCAATGATGATCTCCCTGACATATGCAAA	1380	Db	2401	TCATTA 2406	
Db	1321	GAGTTTGTGTTGAGAAATTCGAATCTGATGATCAATGATGATCTCCCTGACATATGCAAA	1380				
Qy	1381	TTGTTGAAACCTGCTGTCAGAGGCTGCTTTCTTCTAGGTTCAAGACACCAAGATTA 1440					

RESULT 2

US-10-600-070-9
 ; Sequence 9, Application US/10600070
 ; Publication No. US20040139500A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Oeteryoung, Katherine W.
 ; APPLICANT: Vittha, Stanislaw
 ; APPLICANT: Kokeharova, Olga A.
 ; APPLICANT: Gao, Hongo
 ; TITLE OF INVENTION: Plasmid Division and Related Genes and Proteins, and Methods of
 ; TITLE OF INVENTION: Use
 ; FILE REFERENCE: MSU-08153
 ; CURRENT APPLICATION NUMBER: US/10/600,070
 ; NUMBER OF SEQ ID NOS: 206
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 9
 ; LENGTH: 2406
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 US-10-600-070-9

Query March 99.9%; Score 2402.8; DB 7; Length 2406;
 Best Local Similarity 99.9%; Pred No. 0;
 Matches 2404; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 ATGGAAGCTCTGAGTCACGTCGGCATTTGGTCTCTCCCAATTCGAATTAAGCCATTACCA 60
 DB 1 ATGGAAGCTCTGAGTCACGTCGGCATTTGGTCTCTCCCAATTCGAATTAAGCCATTACCA 60
 OY 61 CCGGGAAGCAAAAGCTCCGAGTACCAACACTCTACAACTATCTGCTCCGACG 120
 DB 61 CCGGGAAGCAAAAGCTCCGAGTACCAACACTCTACAACTATCTGCTCCGACG 120
 OY 121 AAATGGCGGACCGCTCTCTCCGACTTCAATTTCACTCCGATTCCTCCCTCTCTCC 180
 DB 121 AAATGGCGGACCGCTCTCTCCGACTTCAATTTCACTCCGATTCCTCCCTCTCTCC 180
 OY 181 TTCCGCAACCGCCACCAACCGCCACTCTCGTCTCTGCAACATTAATGATGTCCTC 240
 DB 181 TTCCGCAACCGCCACCAACCGCCACTCTCGTCTCTGCAACATTAATGATGTCCTC 240
 OY 241 GAAGCCACGTCGCCATCCCATTTGATTTCTACCGAGTATTAAGAGCTCAAAACATTTG 300
 DB 241 GAAGCCACGTCGCCATCCCATTTGATTTCTACCGAGTATTAAGAGCTCAAAACATTTG 300
 OY 301 TTAACCGATGGAATCAGAAAGACATTCGAAGCTAGGGTTTGAACCGCGCAATTCGGT 360
 DB 301 TTAACCGATGGAATCAGAAAGACATTCGAAGCTAGGGTTTGAACCGCGCAATTCGGT 360
 OY 361 TTACGCGACGACGCTTTATATCAGCCGAGACAGATTTCTTCAAGCTGCTTGCAGAACTCTG 420
 DB 361 TTACGCGACGACGCTTTATATCAGCCGAGACAGATTTCTTCAAGCTGCTTGCAGAACTCTG 420
 OY 421 TCTAATCTCGGTCTAGAAAGAGATGACATGAAGGCTCTTGAATGATGAAGAGCTACA 480
 DB 421 TCTAATCTCGGTCTAGAAAGAGATGACATGAAGGCTCTTGAATGATGAAGAGAGCTACA 480
 OY 481 GTCATCATGATGTTCTTGGGATAAGGTTCTGGGGCTCTCTGTATGATGAAGAAGT 540
 DB 481 GTCATCATGATGTTCTTGGGATAAGGTTCTGGGGCTCTCTGTATGATGAAGAAGT 540
 OY 541 GGTGAGACTGAGATGTTCTTGGGTTGATGAGCTCTGCTTAAAGAGAGTTTGCCTAAG 600
 DB 541 GGTGAGACTGAGATGTTCTTGGGTTGATGAGCTCTGCTTAAAGAGAGTTTGCCTAAG 600
 OY 601 TCGTTAAAGCAAGATGTTTATGATATGAGGCTTGGCTTCTGATGTCGAGGGAT 660
 DB 601 TCGTTAAAGCAAGATGTTTATGATATGAGGCTTGGCTTCTGATGTCGAGGGAT 660
 OY 661 GCTATGCAATGATGCACTGATTTTATTAAGTTATGATGATGATGATGATGATGAT 720
 DB 661 GCTATGCAATGATGCACTGATTTTATTAAGTTATGATGATGATGATGATGATGAT 720

OY 721 AAGCTTTTACAGAGAGAGCAAGTATGCTTGCACCGGATTTTACGTCACAAATGAT 780
 DB 721 AAGCTTTTACAGAGAGAGCAAGTATGCTTGCACCGGATTTTACGTCACAAATGAT 780
 OY 781 GAGACTTTGAAAGAGATCACTCCGCGTATGCTTGGAGCTACTTGGCTTCCGCTGCT 840
 DB 781 GAGACTTTGAAAGAGATCACTCCGCGTATGCTTGGAGCTACTTGGCTTCCGCTGCT 840
 OY 841 GATGATTAACGTCGCAAAAAGCTAAATGTTTAAAGCGGTGTCGGAAATATTTTGGTCT 900
 DB 841 GATGATTAACGTCGCAAAAAGCTAAATGTTTAAAGCGGTGTCGGAAATATTTTGGTCT 900
 OY 901 GTTGGAGAGGTGAGACATCACTCTTGTGGGGTTTGAACCCGTGAGAAATTAATGAT 960
 DB 901 GTTGGAGAGGTGAGACATCACTCTTGTGGGGTTTGAACCCGTGAGAAATTAATGAT 960
 OY 961 GAGCGTTTTTACAAATGACAGCTGTCGAGAGAGTATCTTTTGTATGATACCCCAAGC 1020
 DB 961 GAGCGTTTTTATGATATACAGCTGTCGAGAGAGTATCTTTTGTATGATACCCCAAGC 1020
 OY 1021 AATATTCAGAGAGTCAATTTGAAGTTTACGAAGTTGCACTTGTGCTTGTGCTCAAGCT 1080
 DB 1021 AATATTCAGAGAGTCAATTTGAAGTTTACGAAGTTGCACTTGTGCTTGTGCTCAAGCT 1080
 OY 1081 TTTATGTTAAGAAACCAACCTTTTACAGATGCTGATTAAGCAATTCAGCAACTTCAAG 1140
 DB 1081 TTTATGTTAAGAAACCAACCTTTTACAGATGCTGATTAAGCAATTCAGCAACTTCAAG 1140
 OY 1141 CAGGCTAAGTAAATGAGTATGAGATTCCTGCGATGTTGATGATGATGATGATGATGAT 1200
 DB 1141 CAGGCTAAGTAAATGAGTATGAGATTCCTGCGATGTTGATGATGATGATGATGATGAT 1200
 OY 1201 GAGATAGACTTCGCTCTAGAAAGGGGACTCTGTGCACTGCTTATTAAGCAAAATGATGAA 1260
 DB 1201 GAGATAGACTTCGCTCTAGAAAGGGGACTCTGTGCACTGCTTATTAAGCAAAATGATGAA 1260
 OY 1261 TGCCGATATGCTTGGGCTTGAACAAGTGAAGATTCACATTAAGCAAAATGATGATG 1320
 DB 1261 TGCCGATATGCTTGGGCTTGAACAAGTGAAGATTCACATTAAGCAAAATGATGATG 1320
 OY 1321 GAGTTGTTTGAAGAAATTCGAATCGTATGATGATGATGATGATGATGATGATGATG 1380
 DB 1321 GAGTTGTTTGAAGAAATTCGAATCGTATGATGATGATGATGATGATGATGATGATG 1380
 OY 1381 TTGTTGGAACCTGCTTGGGCAAGGGTGTCTTCTCTAGTTCAGAGACCAACCAAGATGAA 1440
 DB 1381 TTGTTGGAACCTGCTTGGGCAAGGGTGTCTTCTCTAGTTCAGAGACCAACCAAGATGAA 1440
 OY 1441 AAATTTAACTCGGGGACTACTATGATGATGATGATGATGATGATGATGATGATGATG 1500
 DB 1441 AAATTTAACTCGGGGACTACTATGATGATGATGATGATGATGATGATGATGATGATG 1500
 OY 1501 GAGGTATGTCAGGGTCTCTCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560
 DB 1501 GAGGTATGTCAGGGTCTCTCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560
 OY 1561 CATGTGAAGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATG 1620
 DB 1561 CATGTGAAGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATG 1620
 OY 1621 AGAAACTCGGCTGAACCCAGAGATGTCAGAGACAGTGTATGATGATGATGATGATGATG 1680
 DB 1621 AGAAACTCGGCTGAACCCAGAGATGTCAGAGACAGTGTATGATGATGATGATGATGATG 1680
 OY 1681 AACATGATGAGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1740
 DB 1681 AACATGATGAGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1740
 OY 1741 GAAACTTTGAAACTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATG 1800
 DB 1741 GAAACTTTGAAACTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATG 1800
 OY 1801 GAAACTCTGTTGAATATGTCGCTTGTGATATGTTTAAAGAGGCAAGTGTGAAGATCCTA 1860

1801 GAAATCTGTTGAAATGTCGGTGTGATGTTAAAGGAGGCGAAGTGAAGTCTTA 1860
1861 GCTGCTGGTGGCAATTTGACATGATTTCACTGTTACAGCAGAAATATTTCTTAAAGC 1920
1861 GCTGCTGGTGGCAATTTGACATGATTTCACTGTTACAGCAGAAATATTTCTTAAAGC 1920
1921 AGCTCATCTTTTCAACGCAAGGATATGTTCTTCTATGAAATCTGATGTCGATCA 1980
1921 AGCTCATCTTTTCAACGCAAGGATATGTTCTTCTATGAAATCTGATGTCGATCA 1980
1981 GGGTCAATGCAAGCTGCAAGATTTCAAGACATTTCCCAAGATGATGCTGATGCAAG 2040
1981 GGGTCAATGCAAGCTGCAAGATTTCAAGACATTTCCCAAGATGATGCTGATGCAAG 2040
2041 AATATGATCCCAAGTGGCAGAAATTAAGTCTGCTGCTTTGGGCTGATCAACGCA 2100
2041 AATATGATCCCAAGTGGCAGAAATTAAGTCTGCTGCTTTGGGCTGATCAACGCA 2100
2101 GAAATGTTACCAAGATTTGATGAGGCAATGCTGAAGATTTGATGCAAGACAGCT 2160
2101 GAAATGTTACCAAGATTTGATGAGGCAATGCTGAAGATTTGATGCAAGACAGCT 2160
2161 GAAATGTTACCAAGATTTGATGAGGCAATGCTGAAGATTTGATGCAAGACAGCT 2220
2161 GAAATGTTACCAAGATTTGATGAGGCAATGCTGAAGATTTGATGCAAGACAGCT 2220
2221 GTGACAGTCTCAGCAGATGGAACCCGCTGCTGCTGAGGAAGCACTGAGAGAGCTGCT 2280
2221 GTGACAGTCTCAGCAGATGGAACCCGCTGCTGCTGAGGAAGCACTGAGAGAGCTGCT 2280
2281 TGTCTATCTGATTTGTTGTTTATCTACAGAAAACAATGCTGATGTCAGAACCTTAC 2340
2281 TGTCTATCTGATTTGTTGTTTATCTACAGAAAACAATGCTGATGTCAGAACCTTAC 2340
2341 AGATACGAAGTTTCTGCTCCAGTCAAGGCTGGAAGAAATCACTGAAGGCTCTGTTTGA 2400
2341 AGATACGAAGTTTCTGCTCCAGTCAAGGCTGGAAGAAATCACTGAAGGCTCTGTTTGA 2400
2401 TCATTA 2406
2401 TCATTA 2406

RESULT 3
US-10-600-070-128
; Sequence 128, Application US/10600070
; Publication No. US2004013950A1
; GENERAL INFORMATION:
; APPLICANT: Osteoryoung, Katherine W.
; APPLICANT: Vitina, Stanislaw
; APPLICANT: Kokeharova, Olga A.
; APPLICANT: Gao, Hongbo
; TITLE OF INVENTION: Placitid Division and Related Genes and Proteins, and Methods of
; FILE REFERENCE: MSU-08153
; CURRENT APPLICATION NUMBER: US/10/600,070
; CURRENT FILING DATE: 2003-06-20
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 128
; LENGTH: 2406
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-600-070-128

Query Match 99.7%; Score 2399.6; DB 7; Length 2406;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2402; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
1 ATGGAAGCTCTGAGTCAGTGGGCAATGGTCTCTCCCAATTCGAATTATGCGGATTAACA 60
1 ATGGAAGCTCTGAGTCAGTGGGCAATGGTCTCTCTCCCAATTCGAATTATGCGGATTAACA 60

61 CCGGCGACGACAAAGCTCCGACGATGACCAACACCTTACAACTATCTGCTCCGACG 120
61 CCGGCGACGACAAAGCTCCGACGATGACCAACACCTTACAACTATCTGCTCCGACG 120
121 AATATGCGGACGATCTCTCTCCGATTTCAATTTTCACTCCGATTTCTCTCTCTCT 180
121 AATATGCGGACGATCTCTCTCCGATTTCAATTTTCACTCCGATTTCTCTCTCTCT 180
181 TTGCGCACCGCCACCAACCGCACTCTGCTCTGCGCAACATATGATTCGCTCC 240
181 TTGCGCACCGCCACCAACCGCACTCTGCTCTGCGCAACATATGATTCGCTCC 240
241 GAAAGCGACGCTCCCATCCCATTTGATTTCTACAGGATTTAGAGAGCTCAAAACATTT 300
241 GAAAGCGACGCTCCCATCCCATTTGATTTCTACAGGATTTAGAGAGCTCAAAACATTT 300
301 TTAACCGATGGAATCAGAAAGCAATTTGAGGCTTTCGAAACCGCGCAATTCGCT 360
301 TTAACCGATGGAATCAGAAAGCAATTTGAGGCTTTCGAAACCGCGCAATTCGCT 360
361 TTCAGCAGCAGCGCTTAAATCAGCCGAGACAGATTTCTCAAGCTGCTTCCGAACTCT 420
361 TTCAGCAGCAGCGCTTAAATCAGCCGAGACAGATTTCTCAAGCTGCTTCCGAACTCT 420
421 TCTATCTCCTGCTTACAGAGAGTACATGAAGTCTTCTGATGATGAAAGAGCTTAC 480
421 TCTATCTCCTGCTTACAGAGAGTACATGAAGTCTTCTGATGATGAAAGAGCTTAC 480
481 GTCTATCTGATGCTCTTGGGATTAAGTCTCTGAGGCTCTCTGATGATGAAAGAGT 540
481 GTCTATCTGATGCTCTTGGGATTAAGTCTCTGAGGCTCTCTGATGATGAAAGAGT 540
541 GGTAGACAGTGAATGATCTTCCGCTTGTGAGGCTCTGCTTAAAGAGAGGTTGCCAAG 600
541 GGTAGACAGTGAATGATCTTCCGCTTGTGAGGCTCTGCTTAAAGAGAGGTTGCCAAG 600
601 TCGTTTAAAGCAAGTGTGTTTATGATGATGATGATGATGATGATGATGATGATGAT 660
601 TCGTTTAAAGCAAGTGTGTTTATGATGATGATGATGATGATGATGATGATGATGAT 660
661 GCTATGCGATTTGATTCACCTGATTTTATTAACGTTTATGATGATGATGATGATGAT 720
661 GCTATGCGATTTGATTCACCTGATTTTATTAACGTTTATGATGATGATGATGATGAT 720
721 AAGCTTTTACAGAGAGAGCAAGTACCTTGCACCGGATTTACGTCGACAAATTTGAT 780
721 AAGCTTTTACAGAGAGAGCAAGTACCTTGCACCGGATTTACGTCGACAAATTTGAT 780
781 GAGACTTTGGAAGAGATCACTCCGCTTATGCTTGGAGCTTACCTGCTTACCGCTTGT 840
781 GAGACTTTGGAAGAGATCACTCCGCTTATGCTTGGAGCTTACCTGCTTACCGCTTGT 840
841 GATATATTCGCTGGAAGAAAGCTTAAATGTTTAAAGCGTGTGCGGAATTTTGTGCT 900
841 GATATATTCGCTGGAAGAAAGCTTAAATGTTTAAAGCGTGTGCGGAATTTTGTGCT 900
901 GTTGAAGAGGTGAGATCAGCTCTTGTGGGGGTTTGAACCGTGAGAAATTTATGAT 960
901 GTTGAAGAGGTGAGATCAGCTCTTGTGGGGGTTTGAACCGTGAGAAATTTATGAT 960
961 GAGGCGTTTTTACGAATGACAGCTGCTGAGCAGGTTGATCTTTTGTAGCTAACCCAGC 1020
961 GAGGCGTTTTTACGAATGACAGCTGCTGAGCAGGTTGATCTTTTGTAGCTAACCCAGC 1020
1021 AATATTTACAGAGAGTATTTGAAGTTTAAAGCTTGAAGCTTCTGAGCTCAAGCT 1080
1021 AATATTTACAGAGAGTATTTGAAGTTTAAAGCTTGAAGCTTCTGAGCTCAAGCT 1080
1081 TTTATGTTGAAGAGCAGCCTTTTACAGGATGCTATTAAGCAATTTCCAGCAATTTCA 1140
1081 TTTATGTTGAAGAGCAGCCTTTTACAGGATGCTATTAAGCAATTTCCAGCAATTTCA 1140

1141 CAGGCTAAGCTAATGCTATGAGATTCTCGAGTCTGTATGATACACGAAATATGCG 1200
1141 CAGGCTAAGCTAATGCTAATGAGATCTCGAGTCTGTATGATACACGAAATATGCG 1200
1201 GAGATAGACTTGGCTAGAAAAGGGGACCTGTGCACTGCTTAATAGGCAAACTGATGGA 1260
1201 GAGATAGACTTGGCTAGAAAAGGGGACCTGTGCACTGCTTAATAGGCAAACTGATGGA 1260
1261 TCCCGATATGCTGCTTGAAGATTCAGATGAGATTCAGATATAGGATTCAGATATGCTG 1320
1261 TCCCGATATGCTGCTTGAAGATTCAGATGAGATTCAGATATAGGATTCAGATATGCTG 1320
1321 GAGTTGTTTGGAGAAATTCATATGATGATGATGATGATGATGATGATGATGATGATG 1380
1321 GAGTTGTTTGGAGAAATTCATATGATGATGATGATGATGATGATGATGATGATGATG 1380
1381 TTGTTGGAACCTGGTGGCAGGGGTTGTTCTTCCAGTTCAGAGACCAAGATTA 1440
1381 TTGTTGGAACCTGGTGGCAGGGGTTGTTCTTCCAGTTCAGAGACCAAGATTA 1440
1441 AAATTTAACTCGGGGACTACTATGATGATGATGATGATGATGATGATGATGATGATG 1500
1441 AAATTTAACTCGGGGACTACTATGATGATGATGATGATGATGATGATGATGATGATG 1500
1501 GAGGTAGTTGAGGTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1560
1501 GAGGTAGTTGAGGTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1560
1561 CATGGAAGCTAGTGTCTATGAGGACCTGAGGAAAGTTTCTTCTTCTTCTTCTTCTTCT 1620
1561 CATGGAAGCTAGTGTCTATGAGGACCTGAGGAAAGTTTCTTCTTCTTCTTCTTCTTCT 1620
1621 AGAACTCGGCTGAAACCAAGATGTCAGAGACAGTGTGATGATGATGATGATGATGATG 1680
1621 AGAACTCGGCTGAAACCAAGATGTCAGAGACAGTGTGATGATGATGATGATGATGATG 1680
1681 AACAATGATGAGCCGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
1681 AACAATGATGAGCCGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
1741 GAAACTTTGAACTAATGATTAATGCAATTCAGAGCTGGGCTTCAGAGATGAGCTTGA 1800
1741 GAAACTTTGAACTAATGATTAATGCAATTCAGAGCTGGGCTTCAGAGATGAGCTTGA 1800
1801 GAAACTTTGAACTAATGATTAATGCAATTCAGAGCTGGGCTTCAGAGATGAGCTTGA 1860
1801 GAAACTTTGAACTAATGATTAATGCAATTCAGAGCTGGGCTTCAGAGATGAGCTTGA 1860
1861 GCTGCTGGTGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1920
1861 GCTGCTGGTGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1920
1921 AGCTATCTTTTCAACGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1980
1921 AGCTATCTTTTCAACGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1980
1981 GGGTCACTCAAGCTGAGATTCAGAGACCTTCCAGAAATGATGATGATGATGATGATGAT 2040
1981 GGGTCACTCAAGCTGAGATTCAGAGACCTTCCAGAAATGATGATGATGATGATGATGAT 2040
1991 GGGTCACTCAAGCTGAGATTCAGAGACCTTCCAGAAATGATGATGATGATGATGATGAT 2040
1991 GGGTCACTCAAGCTGAGATTCAGAGACCTTCCAGAAATGATGATGATGATGATGATGAT 2040
2041 AATATAGTATCCAGAGCTGAGAGATTAATGATGATGATGATGATGATGATGATGATGAT 2100
2041 AATATAGTATCCAGAGCTGAGAGATTAATGATGATGATGATGATGATGATGATGATGAT 2100
2101 GAAATGTTTACCAAGAGCTTGGATGGGGAATGATGATGATGATGATGATGATGATGATGAT 2160
2101 GAAATGTTTACCAAGAGCTTGGATGGGGAATGATGATGATGATGATGATGATGATGATGAT 2160
2161 GAAACTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 2220
2161 GAAACTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 2220
2221 GTGACAGTCTCAGCAGATGAAACCCGTGCTCTGCTGAGGAACTTCTGAGAGATCTTCT 2280

2221 GTGACAGTCTCAGCAGATGAAACCCGTGCTCTGCTGAGGAACTTCTGAGAGATCTTCT 2280
2281 TGTCTATCTGATTTGGTTCATCCAGAAACAAATGCTACTGATGATGATGATGATGATGAT 2340
2281 TGTCTATCTGATTTGGTTCATCCAGAAACAAATGCTACTGATGATGATGATGATGATGAT 2340
2341 AGATACGAAGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2400
2341 AGATACGAAGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2400
2401 TCATTA 2406
2401 TCATTA 2406
RESULT 4
US-10-600-070-130
; Sequence 130, Application US/10600070
; Publication No. US20040139500A1
; GENERAL INFORMATION:
; APPLICANT: Osteoryoung, Katherine W.
; APPLICANT: Vitsha, Stanislav
; APPLICANT: Koksharova, Olga A.
; APPLICANT: Gao, Hong
; TITLE OF INVENTION: Placid Division and Related Genes and Proteins, and Methods of
; FILE REFERENCE: MSU-08153
; CURRENT APPLICATION NUMBER: US/10/600,070
; CURRENT FILING DATE: 2003-06-20
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 130
; LENGTH: 2637
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-600-070-130
Query Match 99.7%; Score 2399.6; DB 7; Length 2637;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2402; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
1 ATGGAAGCTCTGAGTCAAGTGGGATTTGCTCTCCCAATTCATTAATGCGATTAACCA 60
114 ATGGAAGCTCTGAGTCAAGTGGGATTTGCTCTCCCAATTCATTAATGCGATTAACCA 173
61 CCGGCGAGCAAAAGCTCCGACGTAAGCCAAACACCTTAACTATCTGCTCCGACAGC 120
174 CCGGCGAGCAAAAGCTCCGACGTAAGCCAAACACCTTAACTATCTGCTCCGACAGC 233
121 AATGAGCGAGCGCTCTCTCCGACTTCAATTTACCTCGATTCCTCCCTCCCTCC 180
234 AATGAGCGAGCGCTCTCTCCGACTTCAATTTACCTCGATTCCTCCCTCCCTCC 293
181 TTGCGCAACCGCACACACCGCACCTCTGCTCTCTGCGACCATTAATGATGCTGCC 240
294 TTGCGCAACCGCACACACCGCACCTCTGCTCTCTGCGACCATTAATGATGCTGCC 353
241 GAAAGCGACGCTCCCATCCCATTTTCAACAGGATTAATGAGACTCAACACATTTT 300
354 GAAAGCGACGCTCCCATCCCATTTTCAACAGGATTAATGAGACTCAACACATTTT 413
301 TTAACCGATGATCAGAAAGATTCGAGAGCTAGGTTTGAACCGCGCAATTCGGT 360
414 TTAACCGATGATCAGAAAGATTCGAGAGCTAGGTTTGAACCGCGCAATTCGGT 473
361 TTACGCGACGCTTTTATATCAGCGGAGACAGATTTCTTCAAGCTGCTTGGAACCTCTG 420
474 TTACGCGACGCTTTTATATCAGCGGAGACAGATTTCTTCAAGCTGCTTGGAACCTCTG 533
421 TCTATATCTCGGTCTAGAAAGAGATTAATGAGAGCTTCTTGAATGATGAAAGAGCTACA 480
534 TCTATATCTCGGTCTAGAAAGAGATTAATGAGAGCTTCTTGAATGATGAAAGAGCTACA 593

QY 481 GTCATCAGTAGTTCCTTGGGATTAAGGTTCCGAGGCTCTCTGTGTATTTGCAAGAGGT 540
Db 594 GTCACTCATAGTTCCTTGGGATTAAGGTTCCGAGGCTCTCTGTGTATTTGCAAGAGGT 653
QY 541 GGTGAGACTGAGATGTTCTTCGGGTTGGTGAAGCTCTGCTTAAGAGAGGTTCCTAAG 600
Db 654 GGTGAGACTGAGATGTTCTTCGGGTTGGTGAAGCTCTGCTTAAGAGAGGTTCCTAAG 713
QY 601 TCGTTTAAGCAAGATGTTAGTTATGCGGCTTGGCTTCTGATGTCTGAGGGAT 660
Db 714 TCGTTTAAGCAAGATGTTAGTTATGCGGCTTGGCTTCTGATGTCTGAGGGAT 773
QY 661 GCTATGCGATTGGATCCACTGATTTTATTACTGGTTATGAGTTTGTGAGAAAGCTTTG 720
Db 774 GCTATGCGATTGGATCCACTGATTTTATTACTGGTTATGAGTTTGTGAGAAAGCTTTG 833
QY 721 AAGCTTTTAACAGAGAGAGAGCAAGTACCTTGAACCGGATTTACGTGCACAAATTTAT 780
Db 834 AAGCTTTTAACAGAGAGAGAGCAAGTACCTTGAACCGGATTTACGTGCACAAATTTAT 893
QY 781 GAGACTTTGAAGAGATCACTCCGGCTTATGTCTTGAAGCTACTTGGCTTACCGCTTGGT 840
Db 894 GAGACTTTGAAGAGATCACTCCGGCTTATGTCTTGAAGCTACTTGGCTTACCGCTTGGT 953
QY 841 GATGATTAACGTCGCAAAAAGCTAAATGTTTAAGCGGTGTGCGGAATATTTTGTGTCT 900
Db 954 GATGATTAACGTCGCAAAAAGCTAAATGTTTAAGCGGTGTGCGGAATATTTTGTGTCT 1013
QY 901 GTTGAAGAGAGGTGAGAGCATCAAGCTCTTGTGGGGTTTGAACCGGATTTATGAT 960
Db 1014 GTTGAAGAGAGGTGAGAGCATCAAGCTCTTGTGGGGTTTGAACCGGATTTATGAT 1073
QY 961 GAGGCGTTTTCAGATGACAGCTGCTGAGCAGGTTGATCTTTTGTGATACCCCAAGC 1020
Db 1074 GAGGCGTTTTCAGATGACAGCTGCTGAGCAGGTTGATCTTTTGTGATACCCCAAGC 1133
QY 1021 AATATTCAGACAGACTCAATTTGAAGTTTACGAAGTTGACATTTGCTCTTGTGGCTCAAGCT 1080
Db 1134 AATATTCAGACAGACTCAATTTGAAGTTTACGAAGTTGACATTTGCTCTTGTGGCTCAAGCT 1193
QY 1081 TTTTATGTTAAGAAAGCCACCTTTTACAGAGTGTGTATTAAGCAATTCACACACTTGAAG 1140
Db 1194 TTTTATGTTAAGAAAGCCACCTTTTACAGAGTGTGTATTAAGCAATTCACACACTTGAAG 1253
QY 1141 CAGGCTAAGTAAATGAGTGAAGATTCCTGAGATGTTGATGATACACGGAATTAATTGG 1200
Db 1254 CAGGCTAAGTAAATGAGTGAAGATTCCTGAGATGTTGATGATACACGGAATTAATTGG 1313
QY 1201 GAGATTAAGCTTCGCTCTAAGAAAGGGGACTCTGTGCACTGCTTATTAAGCAAGTTGATGAA 1260
Db 1314 GAGATTAAGCTTCGCTCTAAGAAAGGGGACTCTGTGCACTGCTTATTAAGCAAGTTGATGAA 1373
QY 1261 TGCCTGATATGTTGGGCTTGAACAGTGAAGATTCACAAATTAAGCAATTCACAGTATTTG 1320
Db 1374 TGCCTGATATGTTGGGCTTGAACAGTGAAGATTCACAAATTAAGCAATTCACAGTATTTG 1433
QY 1321 GAGTTTGTGTTGAGAAATTCGAATTCGATGACAAATGATGATCTCCCTGAGCTATGCAAA 1380
Db 1434 GAGTTTGTGTTGAGAAATTCGAATTCGATGACAAATGATGATCTCCCTGAGCTATGCAAA 1493
QY 1381 TTGTTGAAAACCTGTTGGAGGGGTTGTCTTTCTAGGTTCAAGAACCAACCAAGATTA 1440
Db 1494 TTGTTGAAAACCTGTTGGAGGGGTTGTCTTTCTAGGTTCAAGAACCAACCAAGATTA 1553
QY 1441 AAAATTAACCTCGGGGAGTACTATGATATGATGTTTGAAGTACTTGGAAAGAGTG 1500
Db 1554 AAAATTAACCTCGGGGAGTACTATGATATGATGTTTGAAGTACTTGGAAAGAGTG 1613
QY 1501 GAGGTAGTTCAGGGGTTCTCTTGTAGCTCTGCTGCAATATGAGCAAGGATTTGAGCCGAG 1560
Db 1614 GAGGTAGTTCAGGGGTTCTCTTGTAGCTCTGCTGCAATATGAGCAAGGATTTGAGCCGAG 1673

QY 1561 CATGTGAAGCTAGTGTATGACAGCACTGACAGAAAGTTTTCCTTCCCGTATACAGAT 1620
Db 1674 CATGTGAAGCTAGTGTATGACAGCACTGACAGAAAGTTTTCCTTCCCGTATACAGAT 1733
QY 1621 AGAACTCGGTGAACCCAGAGATGTCAGAGACATGTTTACTGTATGATCCCTTGGT 1680
Db 1734 AGAACTCGGTGAACCCAGAGATGTCAGAGACATGTTTACTGTATGATCCCTTGGT 1793
QY 1681 AACATGTAGCCCTGATGAGCTGAGTGTCTTTATTTGACAGAGCTGTAAAGCCCTCT 1740
Db 1794 AACATGTAGCCCTGATGAGCTGAGTGTCTTTATTTGACAGAGCTGTAAAGCCCTCT 1853
QY 1741 GAAAACCTTGAACCTAATGATTAATGCAATTCGACCTGGGGTCTCAGAGAGTACGTTGAT 1800
Db 1854 GAAAACCTTGAACCTAATGATTAATGCAATTCGACCTGGGGTCTCAGAGAGTACGTTGAT 1913
QY 1801 GAAACTACTGTGAAATGTCCTGTTGCTGATATGTTAAAGAGGCAAGTGTGAAGATCTTA 1860
Db 1914 GAAACTACTGTGAAATGTCCTGTTGCTGATATGTTAAAGAGGCAAGTGTGAAGATCTTA 1973
QY 1861 GCTGCTGGTGGCAATTTGACATGTTCACTGTTCAAGCCAGAGATATTTTCTTAAAGC 1920
Db 1974 GCTGCTGGTGGCAATTTGACATGTTCACTGTTCAAGCCAGAGATATTTTCTTAAAGC 2033
QY 1921 AGCTCATCTTTTCAAGCCAGAGATATGTTTCTTATGAAATCTGATGTGCTTACCATA 1980
Db 2034 AGCTCATCTTTTCAAGCCAGAGATATGTTTCTTATGAAATCTGATGTGCTTACCATA 2093
QY 1981 GGGTCAGTCAGAGCTGACAGATTCAGAGCACTTCCAGAAATGATGCTAGAGCTGCAGAG 2040
Db 2094 GGGTCAGTCAGAGCTGACAGATTCAGAGCACTTCCAGAAATGATGCTAGAGCTGCAGAG 2153
QY 2041 AATATGATTAACAGTGGCAGAGATTAAGTCTCTGCTTTTGGGCTGATCAACCGATA 2100
Db 2154 AATATGATTAACAGTGGCAGAGATTAAGTCTCTGCTTTTGGGCTGATCAACCGATA 2213
QY 2101 GAAATGTTTACCAAGGTTTGGATGGCGAATGCTGAAGATTTGAGTGAACAGCACT 2160
Db 2214 GAAATGTTTACCAAGGTTTGGATGGCGAATGCTGAAGATTTGAGTGAACAGCACT 2273
QY 2161 GAAACCTGGCAGCTGGGTTGTTTATGATTAACACTGTGAAACCTATCTGTTGAACGT 2220
Db 2274 GAAACCTGGCAGCTGGGTTGTTTATGATTAACACTGTGAAACCTATCTGTTGAACGT 2333
QY 2221 GTGACAGTCTCAGAGATGGAACCCGCTCTGCTGGTGAACCAACTCTGAGAGAGTCTGCT 2280
Db 2334 GTGACAGTCTCAGAGATGGAACCCGCTCTGCTGGTGAACCAACTCTGAGAGAGTCTGCT 2393
QY 2281 TGTCTATCTGATTTGTTCAATCCAGAAAACAATGCTATGATGTCAGAACTTACACACA 2340
Db 2394 TGTCTATCTGATTTGTTCAATCCAGAAAACAATGCTATGATGTCAGAACTTACACACA 2453
QY 2341 AGATACGAAGTTTCTGCTCCAGAGTCAAGGCTGAGAAAATCACTGAAGGCTCTGTTCTTGA 2400
Db 2454 AGATACGAAGTTTCTGCTCCAGAGTCAAGGCTGAGAAAATCACTGAAGGCTCTGTTCTTGA 2513
QY 2401 TCATTA 2406
Db 2514 TCATTA 2519

RESULT 5
US-10-739-930-227
; Sequence 227, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; FILE REFERENCE: 38-21(53377)B
; CURRENT APPLICATION NUMBER: US/10/739, 930
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 11088

```
; SEQ ID NO 227
; LENGTH: 2679
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: Clone ID: ARATH-23APR03-CLUSTER13643_1
; US-10-739-930-227

Query Match      99.7%; Score 2399.6; DB 8; Length 2679;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2402; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGGAAGCTCTGAGTCACGTCGCAATGCTCTCTCCCAATCCAAATTATGCGGATTACCA 60
DB 115 ATGGAAGCTCTGAGTCACGTCGCAATGCTCTCTCCCAATCCAAATTATGCGGATTACCA 174
QY 61 CCGGGAGACGCAAAAGCTCCGAGTACGCAACACCTCTACAACTATCTGCTCCGACG 120
DB 175 CCGGGAGACGCAAAAGCTCCGAGTACGCAACACCTCTACAACTATCTGCTCCGACG 234
QY 121 AAATGGGCGGACCGCTCTCTCCGACTTCAATTTCACTCCGATTCCTCTCTCTCC 180
DB 235 AAATGGGCGGACCGCTCTCTCTCCGACTTCAATTTCACTCCGATTCCTCTCTCTCC 294
QY 181 TTGCGCACCGGACCAACACCGGCACTCTGCTCTCTGCAACATTAATGATGCTCC 240
DB 295 TTGCGCACCGGACCAACACCGGCACTCTGCTCTCTGCAACATTAATGATGCTCC 354
QY 241 GAAGGCGACGTCGCCATCCCATTTGATTTCTACAGAGTATGAGGCTTCAACATTTG 300
DB 355 GAAGGCGACGTCGCCATCCCATTTGATTTCTACAGAGTATGAGGCTTCAACATTTG 414
QY 301 TTAAACGATGGAATGAGAAGACATTCGAAAGTAGGGTTTCGAAACCGGCGCAATTCGGT 360
DB 415 TTAAACGATGGAATGAGAAGACATTCGAAAGTAGGGTTTCGAAACCGGCGCAATTCGGT 474
QY 361 TTCAAGGACGACGCTTTTAATCAAGCGGAGACAGATTTCTTCAAGCTGCTTTCGAACTCTG 420
DB 475 TTCAAGGACGACGCTTTTAATCAAGCGGAGACAGATTTCTTCAAGCTGCTTTCGAACTCTG 534
QY 421 TCTAATCCCTGGCTTGAAGAAGATGCAATGAAGGCTCTTGAAGATGAAGAAGACTTCA 480
DB 535 TCTAATCCCTGGCTTGAAGAAGATGCAATGAAGGCTCTTGAAGATGAAGAAGACTTCA 594
QY 481 GTCACTACGTATGTTCTTTGGGATTAAGTTCTTGGGGCTCTCTGTATTTGCAAGAAGT 540
DB 595 GTCACTACGTATGTTCTTTGGGATTAAGTTCTTGGGGCTCTCTGTATTTGCAAGAAGT 654
QY 541 GGTGAGACTGAGATAGTTCTTCGGGTTGGTGAAGGCTCTGCTTAAGAGAAGGTTGCTTAAG 600
DB 655 GGTGAGACTGAGATAGTTCTTCGGGTTGGTGAAGGCTCTGCTTAAGAGAAGGTTGCTTAAG 714
QY 601 TCGTTAAGCAAGATGCTTTTAATTAATGAGGCTTGGCTTCTGAGATCTCCGAGGAT 660
DB 715 TCGTTAAGCAAGATGCTTTTAATTAATGAGGCTTGGCTTCTGAGATCTCCGAGGAT 774
QY 661 GCTATGCAATGATTCACCTGATTTTATTAAGTGTATGAGTTTGTGAGGAAGCTTTG 720
DB 775 GCTATGCAATGATTCACCTGATTTTATTAAGTGTATGAGTTTGTGAGGAAGCTTTG 834
QY 721 AAGCTTTTACAGAGGAAGAGCAAGATAGCTTTCGACCGGATTTACGTGCACAAAATTGAT 780
DB 835 AAGCTTTTACAGAGGAAGAGCAAGATAGCTTTCGACCGGATTTACGTGCACAAAATTGAT 894
QY 781 GAGACTTTGGAAGAGATCACTCCGCTTATGCTTGGAGCTACTTGGCTTACCGCTTGGT 840
DB 895 GAGACTTTGGAAGAGATCACTCCGCTTATGCTTGGAGCTACTTGGCTTACCGCTTGGT 954
QY 841 GATGATTAACGCTGCGCAAAAGACTAAATGATTTAAGCGGTGTGCGGAATATTTTGTGTCT 900
DB 955 GATGATTAACGCTGCGCAAAAGACTAAATGATTTAAGCGGTGTGCGGAATATTTTGTGTCT 1014
QY 901 GTTGAAGAGGTGAGACATCAAGCTTTGTGTGGGGGTTTGACCCGCTGAGAAGATTTATGAT 960
```

```
DB 1015 GTTGAAGAGGTGAGACATCAAGCTTTGTGTGGGGGTTTGACCCGCTGAGAAGATTTATGAT 1074
QY 961 GAGGCGTTTTTACGAATACAGCTGCTGAGACAGGTTGATCTTTTGTAGCTTACCCCAAGC 1020
DB 1075 GAGGCGTTTTTACGAATACAGCTGCTGAGACAGGTTGATCTTTTGTAGCTTACCCCAAGC 1134
QY 1021 AATATTCAGACAGATCAATTTGAAGTTTACGAAGTTGACCTTGTGCTCAAGCT 1080
DB 1135 AATATTCAGACAGATCAATTTGAAGTTTACGAAGTTGACCTTGTGCTCAAGCT 1194
QY 1081 TTTATTTGTAAGAAACCAACCTTTTACAGATGCTGATTAAGCAATTCACGAATTCAG 1140
DB 1195 TTTATTTGTAAGAAACCAACCTTTTACAGATGCTGATTAAGCAATTCACGAATTCAG 1254
QY 1141 CAGGCTAAGTAAATGAGCTATGAGATTCCTGCGATGTTGTATGATTAACCGGAATTAATGG 1200
DB 1255 CAGGCTAAGTAAATGAGCTATGAGATTCCTGCGATGTTGTATGATTAACCGGAATTAATGG 1314
QY 1201 GAGATAGACTTCGGCTTGAAGAAAGGGAATCTGTGCACTGCTTATAGGCAAGTTGATGA 1260
DB 1315 GAGATAGACTTCGGCTTGAAGAAAGGGAATCTGTGCACTGCTTATAGGCAAGTTGATGA 1374
QY 1261 TTCCGATGCTGCTTGAAGCAAGTGAAGATTCACAAATTAAGGAATTCAGCTATTGTG 1320
DB 1375 TTCCGATGCTGCTTGAAGCAAGTGAAGATTCACAAATTAAGGAATTCAGCTATTGTG 1434
QY 1321 GAGTTGTTTTTGGGAATTCAAATCGATGATGACATATGATATCCCTGCGACTATGCAAA 1380
DB 1435 GAGTTGTTTTTGGGAATTCAAATCGATGATGACATATGATATCCCTGCGACTATGCAAA 1494
QY 1381 TTGTGGAACCACTGCTTGGCAAGGGGTTGCTTCTCTAGGTTCAAGACACCAAAAGATTA 1440
DB 1495 TTGTGGAACCACTGCTTGGCAAGGGGTTGCTTCTCTAGGTTCAAGACACCAAAAGATTA 1554
QY 1441 AAATTTAAACTCGGGGACTACTATGATGATCTATGTTTGAATTAATTGAAAAGATG 1500
DB 1555 AAATTTAAACTCGGGGACTACTATGATGATCTATGTTTGAATTAATTGAAAAGATG 1614
QY 1501 GAGGTAGTCAAGGTTTCTCTTATGCTGCTGCTGCAACTATGGAAGATTTGGAGCCGAG 1560
DB 1615 GAGGTAGTCAAGGTTTCTCTTATGCTGCTGCTGCAACTATGGAAGATTTGGAGCCGAG 1674
QY 1561 CATGTGAAGACTATGCTATGCAAGCACTGCAAGAAAGTTTCTTCCCGCTATACAGAT 1620
DB 1675 CATGTGAAGACTATGCTATGCAAGCACTGCAAGAAAGTTTCTTCCCGCTATACAGAT 1734
QY 1621 AGAAACTCGGCTGAACCCAAAGATGTGCAAGACAGTGTTAATGATCTGTGTGT 1680
DB 1735 AGAAACTCGGCTGAACCCAAAGATGTGCAAGACAGTGTTAATGATCTGTGTGT 1794
QY 1681 AACATATGAGCCGCTGATGATGAGACCTGCTTATTTGGAAGAGCTGTAAAGCCCTCT 1740
DB 1795 AACATATGAGCCGCTGATGATGAGACCTGCTTATTTGGAAGAGCTGTAAAGCCCTCT 1854
QY 1741 GAAAACTTTGAACTAATGATTAATGCAATTCGACTGGGCTCTGAGAGATGAGTTGAT 1800
DB 1855 GAAAACTTTGAACTAATGATTAATGCAATTCGACTGGGCTCTGAGAGATGAGTTGAT 1914
QY 1801 GAAACTACTGTTGAATATGCTGTTGCTGATATGTTAAAGAGAGCAAGTGTGAAGATCTTA 1860
DB 1915 GAAACTACTGTTGAATATGCTGTTGCTGATATGTTAAAGAGAGCAAGTGTGAAGATCTTA 1974
QY 1861 GCTGCTGCTGCTGCAATTTGACATGCTGCTGCAAGCAAGATTTTCTTAAAGC 1920
DB 1975 GCTGCTGCTGCTGCAATTTGACATGCTGCTGCAAGCAAGATTTTCTTAAAGC 2034
QY 1921 AGCTCATCTTTTCAACCAAGATATGATTTCTTCTATGGAATCTGATGCTGATACATA 1980
DB 2035 AGCTCATCTTTTCAACCAAGATATGATTTCTTCTATGGAATCTGATGCTGATACATA 2094
QY 1981 GGGTCACTCAGAGCTGACGATTCAGAAAGCACTTCCGAATGATGCTTAGAGCTGCAAG 2040
```

Db	2095	GGGTGAGTACGACCTGACGATTCAGAAAGCACTTCCAGAAATGAGTGTCTAGGACTGCGAAG	2154
Qy	2041	AATATAGTATCCAAAGTGGCGAGAGATTAACTCTCTGGCTTTTGGGCTTGATCCGCAATA	2100
Db	2155	AATATAGTATCCAAAGTGGCGAGAGATTAACTCTCTGGCTTTTGGGCTTGATCCGCAATA	2214
Qy	2101	GAAATGTTTCCGAAGGTTTTCGATGGCGGAATGCTGAAGATTGGACTGACAGAGCGAGCT	2160
Db	2215	GAAATGTTTCCGAAGGTTTTCGATGGCGGAATGCTGAAGATTGGACTGACAGAGCGAGCT	2274
Qy	2161	GAACACTGCGCAGCTTGGGTTGTTTATGATTAATACACTGTGTAACTATCTGTGACAGT	2220
Db	2275	GAACACTGCGCAGCTTGGGTTGTTTATGATTAATACACTGTGTAACTATCTGTGACAGT	2334
Qy	2221	GTGACAGCTTTCAGCAATGGAACCCGCTGCTCTGGTGAAGCAATCTTGGAGAGTCTGCT	2280
Db	2335	GTGACAGCTTTCAGCAATGGAACCCGCTGCTCTGGTGAAGCAATCTTGGAGAGTCTGCT	2394
Qy	2281	TGCTATCTGATTTGGTTCATCCAGAAAACAATGCTACTGATGTCAGAACTTACACACA	2340
Db	2395	TGCTATCTGATTTGGTTCATCCAGAAAACAATGCTACTGATGTCAGAACTTACACACA	2454
Qy	2341	AGATACGAAGTTTCTGATCCAAATCGAGGTGAGAAATCACTGAAGGCTCTGTTCTTGCA	2400
Db	2455	AGATACGAAGTTTCTGATCCAAATCGAGGTGAGAAATCACTGAAGGCTCTGTTCTTGCA	2514
Qy	2401	TCATTA 2406	
Db	2515	TCATTA 2520	

QY	241	GAAGCCACGCGCCCATCCCATGATTTTACACAGGTAATTAGAGCTCAAAACATTTG	300
Db	721	GAAGCCACGCGCCCATCCCATGATTTTACACAGGTAATTAGAGCTCAAAACATTTG	780
QY	301	TTAACCGATGGAAATCAGAAAGCATTCGAAAGTGGGTTTCGAAACCGCGCAATTCGT	360
Db	781	TTAACCGATGGAAATCAGAAAGCATTCGAAAGTGGGTTTCGAAACCGCGCAATTCGT	840
QY	361	TTCAAGCGACGCGCTTTAATCAGCCGAGACAGATTCCTCAAGTCTTTGGGAAACTCTG	420
Db	841	TTCAAGCGACGCGCTTTAATCAGCCGAGACAGATTCCTCAAGTCTTTGGGAAACTCTG	900
QY	421	TCTATCCTCGGCTAGAGAAGATCAAGAAGGCTCTCTTGATGATGAAGAAGCTACA	480
Db	901	TCTATCCTCGGCTAGAGAAGATCAAGAAGGCTCTCTTGATGATGAAGAAGCTACA	960
QY	481	GTCAATCACTGATGTTCTTGGGAT-----	504
Db	961	GTCAATCACTGATGTTCTTGGGATTAAGTAATTCGATTCGGAAATAAAGTTTCTTC	1020
QY	505	-----AAGTTCCTGGGGC	518
Db	1021	GTTTAAATTCATGAATYGATTAAGGAAGAACTTTATCTAGTGAAGGTTCTCTGGGGC	1080
QY	519	TCTCTGATATGCAAGAAGGTGGTGAAGACTGAGATAGTCTTTCGGGTGGTGAAGCTCT	578
Db	1081	TCTCTGATATGCAAGAAGGTGGTGAAGACTGAGATAGTCTTTCGGGTGGTGAAGCTCT	1140
QY	579	GCTTAAGAGAGGTGGCTTAAGTGGTTTAAGCAAGATGGGTTTTAGTTATGGCGCTTGC	638

```

, RESULT 6
, US-10-600-070-3
, Sequence 3, Application US/10600070
, Publication No. US20040139500A1
, GENERAL INFORMATION:
, APPLICANT: Oosteryoung, Katherine W.
, APPLICANT: Vittha, Stanislaw
, APPLICANT: Koksharova, Olga A.
, APPLICANT: Gao, Hongbo
, TITLE OF INVENTION: Placid Division and Related Genes and Proteins, and Methods of
, TITLE OF INVENTION: Use
, FILE REFERENCE: MSU-08153
, CURRENT APPLICATION NUMBER: US/10/600,070
, CURRENT FILING DATE: 2003-06-20
, NUMBER OF SEQ ID NOS: 206
, SOFTWARE: PatentIn version 3.2
, SEQ ID NO 3
, LENGTH: 3667
, TYPE: DNA
, ORGANISM: Arabidopsis thaliana
, US-10-600-070-3

```

QY	639	GTTCCTGATGTCCTCGAGGAGTGGTATGAGCATTTGGATCCACTGATTTTATTACTGGTTA	698
Dd	1201	GTTCCTGATGTCCTCGAGGAGTGGTATGAGCATTTGGATCCACTGATTTTATTACTGGTTA	1260
QY	699	TGAGTTTGTTAGAGACCTTTGAAGCTTTTAC-----	730
Dd	1261	TGAGTTTGTTAGAGACCTTTGAAGCTTTTACAGGTAGTTGACTTGCTTGGTAATTTG	1320
QY	731	-----A	731
Dd	1321	ACGAGCGTTGGCTTTATATAGAACTTTCTGATTTTGATACTTTGTTATTAGTCTTGTTGTA	1380
QY	732	GGAGGAAGAGCAAGTAGCTAGCTTGACCGGAAATTACGTGCACAAATTGATGAGCTTTGGA	791
Dd	1381	GGAGGAAGAGCAAGTAGCTAGCTTGACCGGAAATTACGTGCACAAATTGATGAGCTTTGGA	1440
QY	792	AGAGATCACTCCGGCTTATGTCTTGGAGACTACTTGGCTTACCGTTGGTGATGATTATGCG	851
Dd	1441	AGAGATCACTCCGGCTTATGTCTTGGAGACTACTTGGCTTACCGTTGGTGATGATTATGCG	1500

Query Match	71.7%;	Score 1724;	DB 7;	Length 3667;
Best Local Similarity	79.2%;	Pred. No. 0;		
Matches 2406;	Conservative	0;	Mismatches	0;
			Indels	632;
			Gaps	5;

1501 TGCAGAAAGACTAATGCTTTAAGCGGTGTGCGSAAATATTTGTGTGCTCTGTTGAGGAGG 1560

QY	1	TTGGAAGCTCTGAGTCAGTCGGGCAATTGGTCTCTCCCAATTCAAATPAGCCGATTACCA	60
QY	481	ATGGAAGCTCTGAGTCAGTCGGGCAATTGGTCTCTCCCAATTCAAATPAGCCGATTACCA	540
Db	61	CCGGCGAGCAGCAAAAGCTCCGAGTGGACACAACCTCTACAACTATGTGCTCGCGCAGC	120
QY	541	CCGGCGAGCAGCAAAAGCTCCGAGTGGACACAACCTCTACAACTATGTGCTCGCGCAGC	600
Db	121	AAATGGGCGGACCGTCTTCTCTCCGACTTCAAATTTCACTCCGATTCTCTCTCTCTCC	180
QY	601	AAATGGGCGGACCGTCTTCTCTCCGACTTCAAATTTCACTCCGATTCTCTCTCTCTCC	660
Db	181	TTTCGCACCGGCACACACACCGGCCACTCTCTGCTCTCTGCGACCAATCTATTGATGTGTCC	240
QY	661	TTTCGCACCGGCACACACACCGGCCACTCTCTGCTCTCTGCGACCAATCTATTGATGTGTCC	720

Db	1561	TGGAGCAGCAGCTCTTGTTGGGGGTTTGACCCGAGAGATTATGATAGAGCGTTT	1620
QY	972	ACGANTGACAGCTGCTGAGC-----	991
Db	1621	ACGANTGACAGCTGCTGAGCAGATACAGTTAGATACCTTTTAAATTTCTTACGA	1680
QY	992	-----ACGTTGATCTTTTGA	1008
Db	1681	TGATATACCTTTAGGTTCTTCATTTTATGTATGTATGTGTGTGTGATGATCTTTTGA	1740
QY	1009	GCTACCCCAAGCAATATTCACAGCAGAGCATTTGAAGTTTACGAAGTTGCACCTGCTCT	1068
Db	1741	GCTACCCCAAGCAATATTCACAGCAGAGTCATTTGAAGTTTACGAAGTTGCACCTGCTCT	1800

QY 61 CCGGCGACGACAAAGCTCCGAGTACGCCAACACCTCTACAACTATCTGCTCGCGACG 120
DB |||||
QY 541 CCGGCGACGACAAAGCTCCGAGTACGCCAACACCTCTACAACTATCTGCTCGCGACG 600
DB |||||
QY 121 AAATGGGCGGACCGCTTCTCTCCGACTTCGAATTTCACTCCGATTCCTCTCTCTCTCC 180
DB |||||
QY 601 AAATGGGCGGACCGCTTCTCTCCGACTTCGAATTTCACTCCGATTCCTCTCTCTCTCC 660
DB |||||
QY 181 TTGCGCACCGCACACCGACCGCACTCTGCTCTCTGCGACCACTATATGATGATGCC 240
DB |||||
QY 661 TTGCGCACCGCACACCGACCGCACTCTGCTCTCTGCGACCACTATATGATGATGCC 720
DB |||||
QY 241 GAAGCGCAGTCCCATCCCATTTGATTTCTACACGATTTAGAGAGCTCAAAACACATTC 300
DB |||||
QY 301 TTAAACCGATGGAATGGAAGAGCATTTGGAAGTATGAGGTTTGAAAACCGCGCGCATTCG 360
DB |||||
QY 781 TTAAACCGATGGAATGGAAGAGCATTTGGAAGTATGAGGTTTGAAAACCGCGCGCATTCG 840
DB |||||
QY 361 TTCAAGCGACGACGCTTTATCATCGCCGAGACAGATTTCTTCAAGCTGCTTGCGAAAACCTG 420
DB |||||
QY 841 TTCAAGCGACGACGCTTTATCATCGCCGAGACAGATTTCTTCAAGCTGCTTGCGAAAACCTG 900
DB |||||
QY 421 TCTAATCTCGGTTTGAAGAGAGTACAATGAAGGTTCTTGTATGATGAAGAGCTACA 480
DB |||||
QY 901 TCTAATCTCGGTTTGAAGAGAGTACAATGAAGGTTCTTGTATGATGAAGAGCTACA 960
DB |||||
QY 481 GTCATCACTGATGTTCTTGGGAT----- 504
DB |||||
QY 961 GTCATCACTGATGTTCTTGGGATGAAGTAAATTCGATTCGGAATTAATAAGTTTCTTC 1020
DB |||||
QY 505 -----AAGTTTCTCGGAGC 518
DB |||||
QY 1021 GTTTTAATTTGATGAATGGAATGAAGAAAGAACTTTATCTAGTGAAGTTCTCGGAGC 1080
DB |||||
QY 519 TCTCTGTATATTGCAAGAGGTGTGAGACCTGAGATAGTTCTTGGGTTGTGAGGCTCT 578
DB |||||
QY 1081 TCTCTGTATATTGCAAGAGGTGTGAGACCTGAGATAGTTCTTGGGTTGTGAGGCTCT 1140
DB |||||
QY 579 GCTTAAGAGAGGTGCTTAAGTCTTTAAGCAAGATGCTTTTATGATTAAGGCGCTTCG 638
DB |||||
QY 1141 GCTTAAGAGAGGTGCTTAAGTCTTTAAGCAAGATGCTTTTATGATTAAGGCGCTTCG 1200
DB |||||
QY 639 GTTTCCTGATGCTCGAGAGATGCTATGAGCATTTGGATCCACCTGATTTTATTAAGTCTG 698
DB |||||
QY 1201 GTTTCCTGATGCTCGAGAGATGCTATGAGCATTTGGATCCACCTGATTTTATTAAGTCTG 1260
DB |||||
QY 699 TGAAGTTGTTGAAGAGCTTTGAAGCTTTTAC----- 730
DB |||||
QY 1261 TGAAGTTGTTGAAGAGCTTTGAAGCTTTTACAGGTAGTTGACTTGTGTAATTTG 1320
DB |||||
QY 731 -----A 731
DB |||||
QY 1321 ACGAGCGTTGCTTTAATAAGAACTTTCTTGATTTGATACCTTGTATTGAGCTCTTGCTGA 1380
DB |||||
QY 732 GGAGGAAAGAGCAAGTACCTTGTGACCGGATTTTACGTGCACAAATGATGAGACTTTGGA 791
DB |||||
QY 1381 GGAGGAAAGAGCAAGTACCTTGTGACCGGATTTTACGTGCACAAATGATGAGACTTTGGA 1440
DB |||||
QY 792 AGAGATCACTCGCGCTTATGCTCTTGAGACCTTACCTGCTTACCGCTGCTGATGATTAAGC 851
DB |||||
QY 1441 AGAGATCACTCGCGCTTATGCTCTTGAGACCTTACCTGCTTACCGCTGCTGATGATTAAGC 1500
DB |||||
QY 852 TGGGAAAAGACTTAAATGCTTTAAGCGGTGCGGAATTTTGTGCTGTTTGAAGAGG 911
DB |||||
QY 1501 TGGGAAAAGACTTAAATGCTTTAAGCGGTGCGGAATTTTGTGCTGTTTGAAGAGG 1560
DB |||||
QY 912 TGGAGCATCAGCTCTTGTGAGGAGTTTGACCGGTGAGAGTTTATGATGAGGCGTTTTT 971
DB |||||
QY 1561 TGGAGCATCAGCTCTTGTGAGGAGTTTGACCGGTGAGAGTTTATGATGAGGCGTTTTT 1620
DB |||||
QY 972 ACGAATGACAGCTGCTGAGC----- 991
DB |||||

DB |||||
QY 1621 ATGATATCAGCTGCTGAGAGGATATACAGTTTATGATACCTTTTAAATTTCTTTAGCA 1680
DB |||||
QY 992 -----AGSTGATCTTTTGTGA 1008
DB |||||
QY 1681 TGAATTAATCTTAAAGTTTCTCATATTTAATGATGTTGTGTGATGATGATCTTTTGTGA 1740
DB |||||
QY 1009 GCTACCCCAAGCATATATCCAGCAGAGTCAATTTGAAGTTTACGAAGTTGACCTTGTCTT 1068
DB |||||
QY 1741 GCTACCCCAAGCATATATCCAGCAGAGTCAATTTGAAGTTTACGAAGTTGACCTTGTCTT 1800
DB |||||
QY 1069 GTGCTCAAGCTTTTATTTGTTAGAGACCAACCTTTTACAGATGCTGATTAAGCAATTC 1128
DB |||||
QY 1801 GTGCTCAAGCTTTTATTTAGAGAACCAACCTTTTACAGATGCTGATTAAGCAATTC 1860
DB |||||
QY 1129 CAGCACTTCAAGCGGCTAAAGTTAATGAGATGAGATTCCTGATGTTGATGATACA 1188
DB |||||
QY 1861 CAGCACTTCAAGCGGCTAAAGTTAATGAGATGAGATTCCTGATGTTGATGATACA 1920
DB |||||
QY 1189 CGGATTAATTTGGAGATAGACTTCGCTTAGAAAAGGGAAGCTGTGCACTGCTTATAGGC 1248
DB |||||
QY 1921 CGGATTAATTTGGAGATAGACTTCGCTTAGAAAAGGGAAGCTGTGCACTGCTTATAGGC 1980
DB |||||
QY 1249 AAAGTTGATGATGCTGATGATGCTTGGGCTTGAACAGTGAAGATTCACATATAGGAAT 1308
DB |||||
QY 1981 AAAGTTGATGATGCTGATGATGCTTGGGCTTGAACAGTGAAGATTCACATATAGGAAT 2040
DB |||||
QY 1309 CCAGCTATTTGGAGATGTTGTTTGGAGAAATTCAAATGATGATGATGATGATGATGATGAT 1368
DB |||||
QY 2041 CCAGCTATTTGGAGATGTTGTTTGGAGAAATTCAAATGATGATGATGATGATGATGATGAT 2100
DB |||||
QY 1369 GGAATATGCAAAATTTGTGGAAGAACTGTGTGGAAGGAGTTCTTCTTCTAGTTTCAGAGAC 1428
DB |||||
QY 2101 GGAATATGCAAAATTTGTGGAAGAACTGTGTGGAAGGAGTTCTTCTTCTTCTAGTTTCAGAGAC 2160
DB |||||
QY 1429 ACCAAATATTAATAATTTAACTGCGGAGCTACATATGATGATGATGATGATGATGATGAT 1488
DB |||||
QY 2161 ACCAAATATTAATAATTTAACTGCGGAGCTACATATGATGATGATGATGATGATGATGAT 2220
DB |||||
QY 1489 TTGGAAGAGGTGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1548
DB |||||
QY 2221 TTGGAAGAGGTGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2280
DB |||||
QY 1549 ATTGAGCCGAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1608
DB |||||
QY 2281 ATTGAGCCGAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2340
DB |||||
QY 1609 CGCTATACATATGAAACTGCGGCTGAACCCAGAGATGTCAGAGAGAGAGTGTAGTGTGA 1668
DB |||||
QY 2341 CGCTATACATATGAAACTGCGGCTGAACCCAGAGATGTCAGAGAGAGAGTGTAGTGTGA 2400
DB |||||
QY 1669 GATCCTGTGTGATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1728
DB |||||
QY 2401 GATCCTGTGTGATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2460
DB |||||
QY 1729 GTAAAGCCCTCTGAAAACCTTTGAAAACCTAATGATTAATGCAATTTGAGAGCTGAGGAGTCT 1788
DB |||||
QY 2461 GTAAAGCCCTCTGAAAACCTTTGAAAACCTAATGATTAATGCAATTTGAGAGCTGAGGAGTCT 2520
DB |||||
QY 1789 AGTAGCCTTGAATGAAATCTACTGTTGAAATGTCGCTGCTGATATGATTAAGAGGAGCAAGT 1848
DB |||||
QY 2521 AGTAGCCTTGAATGAAATCTACTGTTGAAATGTCGCTGCTGATATGATTAAGAGGAGCAAGT 2580
DB |||||
QY 1849 GTGAAGATCCAGTGTGCTGAGTGTGAGCAATTTGAGACTGATTTCACTGTTCAACCGAAGATAT 1908
DB |||||
QY 2581 GTGAAGATCCAGTGTGCTGAGTGTGAGCAATTTGAGACTGATTTCACTGTTCAACCGAAGATAT 2640
DB |||||
QY 1909 TTTCTTAAAGAGCTCATCTTTTCAACGCAAGATATGATGATGATGATGATGATGATGATGAT 1968
DB |||||
QY 2641 TTTCTTAAAGAGCTCATCTTTTCAACGCAAGATATGATGATGATGATGATGATGATGATGAT 2700
DB |||||
QY 1969 GTGCTTACCA----- 1978
DB |||||

SEQ ID NO 126
LENGTH: 2283
TYPE: DNA
ORGANISM: Oryza sativa
US-10-600-070-126

Query Match 20.0%; Score 481; DB 7; Length 2283;
Best Local Similarity 54.3%; Pred. No. 5,2e-134;
Matches 1256; Conservative 0; Mismatches 880; Indels 177; Gaps 7;

```
Oy 108 CTGCTCCGCCAGCAAAATGGCCGACCGCTCTCTCCGACCTTCAATTCCACTCCGATTC 167
Db 132 CTGGGCGCAAGCGCTCTTCCGCGACTTCACCTCTCCCAACCGCGCGCTCCGACCC 191
Oy 168 CTCCTCTCTCTCTCTGCGCACGCGACCGACCGACCTCTGCTCTGCGACCATC 227
Db 192 GCCGTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTT 251
Oy 228 TATTGATCGTCCCGAAGCGCAAGTCCCATTCCTCCATTTCTACAGGTAATTAGAGC 287
Db 252 CCGCGACGCGCGCGAAGCTCCCTCCGCTCCAGTCGATTTCTACAGGTTCTAGGGGC 311
Oy 288 TCAACAACATTTCTTACCGATGGAATCAAGAGCATTCGAGCTAGGGTTTGAAAC 347
Db 312 AGAGCGCAATTTCTTGGCGATGCGATCAGAGGCGTTGAGGCGACGGAATAGCCACC 371
Oy 348 GCCGCAATTCGGTTTCAAGGACGACGCTTTTATGAGCGGAGACAGATTCTTCAAGCTG 407
Db 372 ACCGCGATATGCTACAGACGAGATGCTCTTGTGTGTGCGCAAAATGCTGCAATTC 431
Oy 408 TTGCGAACTCTGTCTAATCTCGGCTTGAAGAGATGACAATGAAGGCTCTTGTATGA 467
Db 432 CCATGACACTCTCATGAACACAGAACTCCGCACTCAGATGATCGTGGCTTTCTAGAA 491
Oy 468 TGAAGAACTACAGTACATGATGTTCTTGGATTAAGTCTCTGGGGCTCTCTGTGT 527
Db 492 CGGTAAAGGCTCTACCATGATGATTTGTTGGACAAGAGCGCTGG3----- 540
Oy 528 ATTGCAAGAGTGGTGAAGTGAATAGTCTTGGGGTGGTGAAGCTCTGCTTAAGA 587
Db 541 -----GAGGCACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 584
Oy 588 GAGGTGCTTAAGTGTGTTAAGCAAGATGTGGTTTAACTTAAAGGCTTGGCTTTCTGA 647
Db 585 TCGGCAACCAAGCGCTTCAAGCAAGACGTGGTGTACGATGCTGCGCTTATGTGA 644
Oy 648 TGTCTCGAGGATGCTATGCGATTTGCACTGATTTTATTAAGTGTGTGTGTGTGT 707
Db 645 TCTATCAAGGATGCTATGCGACGAAGCCCTCCAGATGTAAATGGCTGCTCGAGTGT 704
Oy 708 TGAAGAACTTTGAAGCTTTTACAGAGAAAGAGCAAGTGAAGCTTGAACCGGATTTAG 767
Db 705 CGAGAGGCTCTCAAGCTCTTGCAGAGAAATGAGCAAGCAATCTCGACCTGATCTCT 764
Oy 768 TGCACAAATTGATGAGACTTTGGAAGAGATCACTCGGCTTATGTCTTGAAGCTTAC 827
Db 765 TTCAAGATTGATGAACCTCTGAGAGATTTACACTCGTGTGTAATTGAGCTTCTCTC 824
Oy 828 CTTACCGCTTGTGATGATTAAGCTGCGAAAGAACTAAATGTTTAAAGCGGTGCGGAA 887
Db 825 CTTTCTTATTAACACAGAGATCATTAAGAGCGCAAGAGGCTTCAAGGTGCGAGAAA 884
Oy 888 TATTTGTGTCGTGTGAGAGAGTGAAGCATCAGCTCTTGTGGGGGTTTGAACCGGGA 947
Db 885 CATTTTGTGAGCGTGTGAGAGAGATGCTACCTGTGAGAGAGATTTTCTCGTGA 944
Oy 948 GAAGTTTATGAATGAGGCGTTTATACGAATGACAGCTGCTGAGCAGGTGATCTTTTGT 1007
Db 945 AGCTTATGAACGAGGCTTTTATGAGATGACATCAATGAACGATGATTTCTTTTC 1004
Oy 1008 AGCTTATGAACGATATTTCCAGAGAGATCATTTGAAGTTTAAAGTTGCACTTGTCT 1067
Db 1005 AAAAAACCGAATGACATTTCTCTGAATGTTGAAAATTTACAAATGTAGACATTGACA 1064
```

```
Oy 1068 TGTGCTCAAGCTTTATTTGTAAGAAGCACACTTTTACAGAGTCTGATTAAGCAATT 1127
Db 1065 TGTGCTCAAGCAATTTAAGTAAAGCCACATTTATCATATGCGGATGATCTTTT 1124
Oy 1128 CCAGCAACTTACAGGCTTAAGGTAATGCTATGAGATTCCTGCGATGTTGATGATAC 1187
Db 1125 TGAACAACCTCAGAAAGTTCAA-----CATAGGTTCTCATTA 1160
Oy 1188 ACGGAATTAATGGAGATTAACCTTCGCTTGAAGGGGACCTGTGCACTGCTTAAG 1247
Db 1161 TGTCTATGATTAATGATGAGACCTTGATTTGAAGGGCATTTCTGCTCATTTGCTAG 1220
Oy 1248 CAAAGTTATGAATGCGGATGATGTTGGCTTGAACAGTGAAGATTTCACAATATAGAA 1307
Db 1221 AGATGTTAGCAAGTGCAGAAATGTGCTTGAATTTGAATTAATGATCTTCAATACAGAA 1280
Oy 1308 TCCAGCTATTGTGAGTTGTTTGGAGAAATTCAAATC---GTATGACAAATGATCT 1364
Db 1281 CCCCCAAATTTCTAGAGTTTATTTGTGACCAACTAGCATGATGAAGAGATGATCTTCT 1340
Oy 1365 CCGTGACTATGCAAAATTTGTGAACCTGTGTCGAGGGGTTGTCTTCTAGGTTGAG 1424
Db 1341 TCCAGGCTGTGCAAGCTTTTGAAGACTTGCTTATCTTTGAGGTTTCTTGAAGACAG 1400
Oy 1425 AGACACCAAGATTTAAATTTTAACTCGGGGACTACTATGATGATCTATGCTTTGAG 1484
Db 1401 AGATACCTCGGGGATGAGATGCTTCAAGCTTGAAGATTAATCAAGATGATCAGAAATTTAAG 1460
Oy 1485 TTAAGTTGAAGAGTGAAGTGAATGATTCAGGGTCTCTTAACTGCTGCTGCAATATGCG 1544
Db 1461 CTACCTTGAAGAGATGAGGAGTGTGTGCTCTCATTTGGCTGCTGCTCTCTCTATTCG 1520
Oy 1545 AAGGATTTGAGCGGAGATGAAAGCTAGTGTCTATGACAGGACTGACAGAAAGTTTTC 1604
Db 1521 AAAAATTGTGTCTCAAGCTACAGTGCATTGGTA----- 1555
Oy 1605 TTCCCGCTATACAGATGAATACTCGGCTGAACCAAGATGTGCAAGAGACAGTGTTCAG 1664
Db 1556 -----CTGTGAATTAAGATTAATTAAGCG----- 1581
Oy 1665 TGTATGCTGTGTTGTAACAATGTAGGCGGTGATGTGTGAGCTGTGTCTTTATTTGACA 1724
Db 1582 -----TTCAACAA 1589
Oy 1725 AGCTGTAAAGCCCTGGAATACTTGAACCTAATGATTAATGCAATTGAGCTGGGCTCTC 1784
Db 1590 GGTTTTCCATTGATTAACAAGTTAGACAGGTCAAGCATGAAATTAATTAAGATGCGCC 1649
Oy 1785 AGAGATGAGGTTGATGAATACTACTGTTGAATGTCGCTGTGATATGTTAAAGAGGC 1844
Db 1650 TGGGGGATATCTTGAATAATTTTGAACAGGAAATGACCTGCTCATGATTCGAGAAATGC 1709
Oy 1845 AAGTGAAGATCTTGAAGTCTGTGTGTGCAATTTGCACTGATTTCACTGTTACGCCAAG 1904
Db 1710 CGCCTTGAATTAATCTCTGCTGCGCATGTTTGTGCACTGTTGCAAGTAAATTTGGGCGCAA 1769
Oy 1905 GTATTTTCTTAAAGCAAGCTCATCTTTTCAACCAAGAGATGATGTTCTTATGGAATC 1964
Db 1770 ATATTGCTCTGTGAGAGGCCCTTCTGTCTATTTAGAGTGAAGATGATCTGTGGAGT 1829
Oy 1965 TGAT-----GTGCTACATAGGCTGATGAGTCAAGTCAAGATTTCAAGACACTTCC 2015
Db 1830 TGTCTAATAGTGTGAGACTCTACTGATGATCTGTGCACTAGATGAAGATCTAATATTC 1889
Oy 2016 CAGAAATGATGCTTGAAGCTGCAAGAAATTAATGATTAATGATGCAAGATTAAGTCTCT 2075
Db 1890 TAGAATGATGCGAAGCTGCGAGAAATTAATGATGCGAAGTGAAGTGAATTAATTA 1949
Oy 2076 GGCCTTTGGGCTGATCAAGCATAGAAATGTTACAGAGATTTTGAATGGGCGAATGCT 2135
Db 1950 GGCCTTGGAGCAAGAACATTTCGTTGATCATTTGCAAGAGGTTCTTGTATGCAACATGCT 2009
```

QY 2136 GAAGATTGGACTGACAGACAGCTGAAACCTGCGAGCTTGGGTTGTTATGATTATAC 2195
DB 2010 AAGAGTGTGACTGACCGGACGAGCGGAGATGAGCGTCAATGGGTGGTTCTGGAGATATAC 2069
QY 2196 ACTGTTGAAACATATCTGTTGACAGTGTGACAGTCTCAGACAGATGAAACCCCTGCTGTGT 2255
DB 2070 ACTATCCGATGAGCATTTGATGATGATCACTATCTCCCTAGATGATGCGAGCGACATGT 2129
QY 2256 GGAAGCAACTGAGAGAGTCTGCTGTATCTGATTTGTTATCCAGAAACATATGC 2315
DB 2130 GGAAGGATGATGATGATGAGGAGCGCAACTTACTATGTTACTGAGCCCAAGAAACATTA 2189
QY 2316 TACTGATGTGACAACTTACAAACAAGATGCAAGATTTCTGCTCCAG--TCAAGGTG 2372
DB 2190 TTTCAATGACACAAATACTACCCGGATGATGATGAGCTTCTCCAACTAGAGAGGTG 2249
QY 2373 GAAATCTGTAAGGCTGTGTTCTTGACATATA 2405
DB 2250 GAAATTAACGGAAGAGAGAGCTCTCAAGTCGTA 2282

RESULT 10
US-10-424-599-129007

; Sequence 129007, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 129007
; LENGTH: 1146
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_87499C.1
US-10-424-599-129007

Query Match 19.7%; Score 472.8; DB 7; Length 1146;
Best Local Similarity 70.5%; Pred. No. 1e-111;
Matches 649; Conservative 0; Mismatches 262; Indels 9; Gaps 1;

QY 107 TCTGCTCCGCGCAAGATGAGCGCAAGCTTCTCTCCGACTTCAATTTCACTCCGAT 166
DB 226 TATCGCCACGCAAAATGGGCGAGCGCTCATTTCCGATTTCCAAATTTCTCGGCGAG 295
QY 167 CCT 226
DB 296 CCGCGCGCTCCACTT-----CCACTCTCACTCAAGCCCCCTCTCCGCTCTCC 346
QY 227 CTATTGATGTCCTCCGACGCGACGCTCCCATCTGATTTCTACAGAGTATTAAGAG 286
DB 347 GCGTGTATCTCTCTGAGCGCTATGCTGTGATCTCTGACTCTGATCCGATTTCTCGGCG 406
QY 287 CTCAAAACATTTCTTAAACGATGATGATGATGATGATGATGATGATGATGATGATG 346
DB 407 CCGAAGCGCACTTCTCGGCGAGCGGCAATTCGAGGCGCTTACGAGCCAAATTTCTGAAAG 466
QY 347 CCGCGCAATTTGCTTCAAGCGAGCGCTTAAATGAGCGGAGAGAGATTTCTTCAAGCTG 406
DB 467 CTCTCAAGAGCGCTTCAAGCGAGCGCTTAAATGAGCGGCGCAAAATTTCTTCAAGAG 526
QY 407 CTGCGAAATCTGCTCTATCTCTCGGTCTGAGAGAGATGATGATGATGATGATGATGATG 466
DB 527 CCGTGGAAACCTTACTGATCT 586
QY 467 ATGAAGAGCTACGATCACTGATGATGATGATGATGATGATGATGATGATGATGATG 526

DB 587 ACGAAGAGCCGCCATTTCTCACTCAAAATCCCTTTCGACAAAGTCTTGAGCGTGTGTG 646
QY 527 TATTGCAAGAGCGGTGTGAGCTGAGATGATGATGATGATGATGATGATGATGATGATGATG 586
DB 647 TGTTCAGAAAGCTGTGAGAGACGAGCTGTGTCTTGAATGTGGAGGAGGATTTGCTTACGG 706
QY 587 AGAGTTCCTTAAGCTGTTTAAAGCAAGATGATGATGATGATGATGATGATGATGATGATGATG 646
DB 707 AGAGTTCCTTAAGCTGTTTAAAGCAAGATGATGATGATGATGATGATGATGATGATGATGATG 766
QY 647 ATGTCTGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 706
DB 767 ACGTGTCAAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 826
QY 707 TTTGAGAGCTTGAAGCTTTTACAG 766
DB 827 TCGAGAGGCGTTGAAGCTTTTGAAG 886
QY 767 GTGCAAAATTTGATGAGACTTTGAGAGAGATCACTCCGCTTATGCTTGAAGCTACTTG 826
DB 887 AAGCAAAATTTGATGAGAGCGCTTGAAGAGATTAACCCAGTTGTGTTTGAAGCTTTAG 946
QY 827 GCTTACGCTTGTGATGATTAACGCTCGGAAAGATTAATGTTTAAAGCGGTGTGCGGA 886
DB 947 CTTGCGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1006
QY 887 ATATTTGTGCTGTTGAGAGAGGTGAGATCAGCTCTGTTGGGGGTTTGAACCGGTG 946
DB 1007 ACATTTGTGCGGTGTGTGTGAGAGGAGGTGAGCAAAATTTGCCGGGTTTCAACCGGTG 1066
QY 947 AGAATTTATGATGAGCGGTTTATACGATGATGATGATGATGATGATGATGATGATGATGATG 1006
DB 1067 AAGCTTCAATGATGAGGATTTCTTACATGATGATGATGATGATGATGATGATGATGATGATG 1126
QY 1007 TAGCTACCCCAAGCAATATT 1026
DB 1127 TAGCCACACCAAGTACTATT 1146

RESULT 11

US-10-425-115-81853
; Sequence 81853, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 81853
; LENGTH: 1411
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_174665C.1
US-10-425-115-81853

Query Match 15.0%; Score 361.4; DB 8; Length 1411;
Best Local Similarity 61.3%; Pred. No. 7.4e-98;
Matches 600; Conservative 0; Mismatches 376; Indels 3; Gaps 1;

QY 96 CTCTCAACTATCTGCTCTCCGCGCAAGATGAGCGGCGGCTCTCTCTCTCTCTCTCTCTCTCT 155
DB 211 CCGCTTCAACCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 270
QY 156 CACCTCGGATTTCT 215

APPLICANT: Gao, Hongo
TITLE OF INVENTION: Placid Division and Related Genes and Proteins, and Methods of
FILE OF INVENTION: Use
FILE REFERENCE: MSU-08153
CURRENT APPLICATION NUMBER: US/10/600,070
CURRENT FILING DATE: 2003-06-20
NUMBER OF SEQ ID NOS: 206
SOFTWARE: PatentIn version 3.2
SEQ ID NO 135
LENGTH: 660
TYPE: DNA
ORGANISM: Medicago truncatula
US-10-600-070-135

Query Match 11.6%; Score 278.8; DB 7; Length 660;
Best Local Similarity 67.3%; Pred. No. 5.2e-73;
Matches 432; Conservative 0; Mismatches 192; Indels 18; Gaps 2;

QY 81 AGCTAGCCACACCTCTACAACTATCTGCTCCGACGAAATGGCCGACCGCTTCT 140
DB 17 ACTTAACCGTCTCCATTCCTCCGCGCGCTCTCCGACACAGTAATGGCGAGACTCAT 76
QY 141 CTCGCACTCAATTCGACGCGATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 200
DB 77 TTCCGATTCCTCAATTCCT 136
QY 201 CGCAGCT 260
DB 137 AGTCACTCTC-----ACTCTCTCTTACCCTCTCTCTCTCTCTCTCTCTCTCTCTCT 190
QY 261 CATTGATTTCTACAGGATTTAGAGCTCAACACATTTCTTACCGATGATGATCAAG 320
DB 191 TCTCGACCTGTCAAAATCTCGCGCGCGAAACGCAATTTCTCGATGATGATTCGAG 250
QY 321 AGCATTCGAAAGTAGGCTTTCGAAACCGCGCAATTCGATTCGAGGAGAGCGCTTAT 380
DB 251 AGCTATGAGAGGAATTCCTCAAGCTCTCTCAATGCTTCAATGATGAGCTTTGAT 310
QY 381 CAGCCGAGACAGATTTCTCAAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 440
DB 311 TGTGCTGCTGAGATTTCTCAAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 370
QY 441 AAGATTCATGAAAGCT 488
DB 371 AAGATTCATGAAAGCT 430
QY 489 TGATGTTCTCTGGAATGAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 548
DB 431 TAAATTCCT 490
QY 549 TGAGATGTTCT 608
DB 491 GAGATGTTCT 550
QY 609 GCAAGATGTTCT 668
DB 551 GCAAGATGTTCT 610
QY 669 ATTGATTCACCTGATTTTATCTAGCTTATGAGATTTGTTGA 710
DB 611 TTGTCTCCGCGCAAGTTTCTATGTTGCTTGTGAGATGCTGA 652

RESULT 14

US-10-424-599-35059
Sequence 35059, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 35059
LENGTH: 1039
TYPE: DNA
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_131660C.1
US-10-424-599-35059

Query Match 9.6%; Score 231.6; DB 7; Length 1039;
Best Local Similarity 64.4%; Pred. No. 1.5e-58;
Matches 380; Conservative 0; Mismatches 204; Indels 6; Gaps 2;

QY 1823 TTGCTGATATGTTAAAGAGGCAAGTGTGAAGATCTTACGCTGCTGTGCGCAATTGAC 1882
DB 50 TTACTGAAGAAATTAACATGCAAGTGTGCAATCATGTGTGTGTGTAATTTGAGAC 109
QY 1883 TGATTTCACTGTTCAGCAGAAATTTCTTAAAGCA---GCTCATCTTTTGAACGA 1939
DB 110 TGGTAACTTTGTTGTTTAAAGTTTAACTTAACTTAAAGATGCTACACCATTTCTCTA 169
QY 1940 AGGATATGTTTCTTCTATGAAATCTGATGCTGCTACCATAGAGGTCAGTCAAGCTGACG 1999
DB 170 AAATGACGTGTTCCGCAATGTGTGATGATGATGATGATGATGATGATGATGATGATG 229
QY 2000 ATTCAAGAG---CACTTCCCAAGATGATGCTAGAGCTGCAAGAGAAATAGTATCAAGT 2056
DB 230 AAAAGTGTAGCAGCTACCAAAATGTAGTGAAGGTTGCCAGGCTTTAATTCGCAAGT 289
QY 2057 GGCAGAAATTAAGTCTGTGCTTTTGGGCGCGATACCGGATGAATGTAATTCACAGAG 2116
DB 230 GGCAGAAATTAAGTCTGTGCTTTTGGGCGCGATACCGGATGAATGTAATTCACAGAG 349
QY 2117 TTTTGGATGAGCGAATGCTGAAATTTGAGCTGACAGAGCTGAAATCGCGAGCTTTG 2176
DB 350 TTTTGGAGCGTGAAGATGTTGAAGATGATGAGACAGATCGTGAAGCTGATGAGAGCGTG 409
QY 2177 GGTGTTTATGATTTATACCTGTTGAACATCTATCTGTTGACAGTGTGACAGTCTGACAG 2236
DB 410 GTTGTCTTATGATCTACCTTGTGAGGACCTCAACATGAGTGTGACATATCACAGA 469
QY 2237 ATGGAACCGTCTGTGTGAGAGCACTGAGAGCTGTGCTGTATCTGATTTGG 2296
DB 470 ATGGCGCTGTGAGTGTGAGAGCACTGAGAGCTGTGATCTGATCTGATGAGCTGAG 529
QY 2297 TTCAATCCAGAAACATGCTATGATGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 2356
DB 530 GTCAATCCGCAACAGATCTCTTAATGACAGACCTGACAGACAGAAAGATGATGATG 589
QY 2357 GGTCAATCTCAGGTTGAGAAATCACTGAGAGCTGTGTTTTCATCATTA 2406
DB 590 TTACAGCGCAGAGTGAATTTGTGAAGAGAGCTGTGATGATCTTAA 639

RESULT 15

US-10-600-070-143
Sequence 143, Application US/10600070
Publication No. US20040139500A1
GENERAL INFORMATION:
APPLICANT: Vitha, Stanislaw
APPLICANT: Koksharova, Olga A.
APPLICANT: Gao, Hongo
TITLE OF INVENTION: Placid Division and Related Genes and Proteins, and Methods of
FILE REFERENCE: MSU-08153
CURRENT APPLICATION NUMBER: US/10/600,070
CURRENT FILING DATE: 2003-06-20
NUMBER OF SEQ ID NOS: 206

; SOFTWARE: Patencin version 3.2
; SEQ ID NO 143
; LENGTH: 537
; TYPE: DNA
; ORGANISM: Trifolium aestivum
US-10-600-070-143

Query Match 9.3%; Score 224; DB 7; Length 537;
Best Local Similarity 63.6%; Pred. No. 1.9e-56;
Matches 341; Conservative 0; Mismatches 195; Indels 0; Gaps 0;

QY 555 AGTTCTTGGGTTGAGGCTCTGCTTAAGAGAGGTTGCTTAAGCTTTAAGCAGA 614
DB 2 AGTGTCTTGCATTTGAGGCGACCTTACTGAGAGCGCCCGCAAGCGGTTCAAGCAGGA 61
QY 615 TGTGGTTTAAATTATGAGGCGCTTGCCTTCTGATGCTGAGAGATGCTATGCAATTGA 674
DB 62 TGTGTGCTGCGCAATGGCGCTGCTTATGTGATCTATCAAGAGAGCGAATGGCGCTAG 121
QY 675 TCCACCTGATTTTATTAAGTTATGAGTTGTGTGAGAGAGCTTTGAAGCTTTTACAGGA 734
DB 122 CCTCCAGATGTAAATCCGCTGCTGTGAGGTGCTTGAAGGCTCTCAAGCTTTTGCAGGA 181
QY 735 GGAAGAGAGATAGCTTGCACCGGATTTACGTGACAAATTGATGAGACTTTGGAAGA 794
DB 182 GGATGGGCAATCAATCTCGACCTGGTTGCTCTCACAAATTGATGAATCTGTGAGGA 241
QY 795 GATCACTCCGCGTTATGCTTGAAGCTTACCGCTTACCGCTTGGATGATTAAGCTGC 854
DB 242 TATCACTCTGCTGTGTGTTTGAAGCTTTTGCCTTCTTGTATGAAAAACATCAGAA 301
QY 855 GAAAAGACTAAATGCTTAAAGCGGTGTGCGAATAATTTGTGCTGTGTGAGAGAGTGG 914
DB 302 TGAACACCAAGAGGCTCTGTGTGAGAAACATTTTGTGAGGTGTGGCAGAGAGG 361
QY 915 AGCATCAGCTCTTGTGTGGGGTTTGAACCGTGAAGATTATGATGAGGCGTTTACG 974
DB 362 TATTGTACTGTGAGAGGATTTTCGGTGAAGCTTACATGAATGAAGCTTCTGCA 421
QY 975 AATGACAGCTGCTGAGAGGTTGATCTTTTGTAGCTACCCCAAGCAATATTCAGCAGA 1034
DB 422 GATGACATGCGCGAGAGATGATTTCTTCAAAAACACCGAATAGCATACCGCTGA 481
QY 1035 GTCAATTTGAAGTTTGAAGTTGCACTTGTGCTTGTGCTCAAGCTTTTATTTGTA 1090
DB 482 ATGTTTGAATCTATAGCGTGGCACTTGCATATGTGCTCAAGCAATTTGTAAGTA 537

Search completed: December 10, 2005, 18:15:31
Job time : 1850.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 10, 2005, 03:43:47 ; Search time 401.33 Seconds
(without alignments)
10656.596 Million cell updates/sec

Title: US-10-600-070B-1

Perfect score: 2406
Sequence: 1 atggaagctctgcagtcacgt.....gctctgtcttcgcatcataa 2406

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/1.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/H.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PTUS.COMB.seq:*
7: /cgn2_6/ptodata/1/ina/PP.COMB.seq:*
8: /cgn2_6/ptodata/1/ina/RE.COMB.seq:*
9: /cgn2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	57.6	2.4	7218	2	US-08-232-463-14
2	50.2	2.1	57280	3	US-09-949-016-11796
3	50.2	2.1	57280	3	US-09-949-016-12843
4	50.2	2.1	57280	3	US-09-949-016-12844
5	50.2	2.1	57280	3	US-09-949-016-12846
6	50.2	2.1	57280	3	US-09-949-016-13542
7	50.2	2.1	57280	3	US-09-949-016-13543
8	50.2	2.1	57280	3	US-09-949-016-13544
9	50.2	2.1	57280	3	US-09-949-016-13545
10	50.2	2.1	57280	3	US-09-949-016-14633
11	50.2	2.1	57280	3	US-09-949-016-14634
12	50.2	2.1	57280	3	US-09-949-016-14635
13	50.2	2.1	57280	3	US-09-949-016-14636
14	50.2	2.1	57280	3	US-09-949-016-14637
15	50.2	2.1	57280	3	US-09-949-016-14638
16	50.2	2.1	57280	3	US-09-949-016-14639
17	50.2	2.1	57280	3	US-09-949-016-14640
18	49.4	2.1	913	2	US-08-217-327-3
19	49.4	2.1	913	2	US-07-885-970A-3
20	49.4	2.1	913	2	US-08-298-687A-3
21	49.4	2.1	913	2	US-08-530-797-2
22	49.4	2.1	913	2	US-08-298-829-3
23	49.4	2.1	913	2	US-08-787-335-2
24	49.4	2.1	1984	2	US-07-885-970A-25

25	49.4	2.1	1985	2	US-08-298-687A-25	Sequence 25, Appl
26	49.4	2.1	1985	2	US-08-298-829-25	Sequence 25, Appl
27	49.2	2.0	324	3	US-09-547-693-234	Sequence 234, Appl
28	48.4	2.0	152132	3	US-09-949-016-13845	Sequence 13845, A
29	48.4	2.0	152145	3	US-09-949-016-12371	Sequence 12371, A
30	46.8	1.9	114793	3	US-10-148-806-3	Sequence 3, Appl1
31	44.8	1.9	832	3	US-09-621-976-2813	Sequence 2813, Ap
32	44.2	1.8	4403765	3	US-09-103-840A-2	Sequence 2, Appl1
33	44.2	1.8	4411529	3	US-09-103-840A-1	Sequence 1, Appl1
34	44.2	1.8	1141	3	US-09-806-708B-22	Sequence 22, Appl1
35	43	1.8	402	3	US-10-131-827-8166	Sequence 8166, Ap
36	43	1.8	1281	3	US-09-991-161-415	Sequence 415, App
37	43	1.8	1281	3	US-09-990-444-415	Sequence 415, App
38	43	1.8	1281	3	US-09-997-333-415	Sequence 415, App
39	43	1.8	1281	3	US-09-992-598-415	Sequence 415, App
40	43	1.8	3265	3	US-09-832-129-13	Sequence 13, Appl
41	42.8	1.8	60990	3	US-09-949-016-14080	Sequence 14080, A
42	42.4	1.8	39154	3	US-09-949-016-12384	Sequence 12384, A
43	42.4	1.8	39154	3	US-09-949-016-12801	Sequence 12801, A
44	42.4	1.8	39443	3	US-09-949-016-14326	Sequence 14326, A
45	42.4	1.8	39443	3	US-09-949-016-14327	Sequence 14327, A

ALIGNMENTS

RESULT 1
US-08-232-463-14
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZgpt-Fls

US-08-232-463-14

Query Match 2.4%; Score 57.6; DB 2; Length 7218;

Best Local Similarity 3.9%; Pred. No. 1.2e-05;

Matches 15; Conservative 218; Mismatches 147; Indels 0; Gaps 0;

QY 30 TCTCTCCCATTCGAATTATGCGAATACACCGCGGACGAAAGCTCCGACGTAGCCA 89

DB 1078 YY 1137

QY 90 CAACACCTCTCAACACTATCTGCTCCGACGAAATGGCCGACCGCTCTCTCTCCGACT 149

DB 1138 YY 1197

QY 150 CAATTTCACCTCCGATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 209

DB 1198 YY 1257

QY 210 CGTCTCTCTGCGACCATCTATGATCGTCCGAAAGCGACGCTCCCATCTCCATGATT 269

DB 1258 YY 1317

QY 270 CTACAGGATATAGAGCTCAACACATTTCTTAACGATGATCAGAAAGCATTCGA 329

DB 1318 YY 1377

QY 330 AGTAGGATTCGAACCGCGCAATTCGTTGACGACGACGCTTATACGCGGAG 389

DB 1378 YY 1437

QY 390 ACAGATCTTCAAGCTGCTT 409

DB 1438 CCAATTCCTCTATCTCTT 1457

RESULT 2

US-09-949-016-11796/c

Sequence 11796, Application US/09949016

Patent No. 6812339

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

CURRENT APPLICATION NUMBER: US/09/949,016

PRIOR FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 11796

LENGTH: 57280

TYPE: DNA

ORGANISM: Human

US-09-949-016-11796

Query Match 2.1%; Score 50.2; DB 3; Length 57280;

Best Local Similarity 53.9%; Pred. No. 0.0059;

Matches 103; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 85 AGCCACAACACCTCTACACATCTGCTCCGCGACGAAATGGCGGACCGCTCTCTCTCC 144

DB 13101 AGCACCATACACACACCAAGCCCTCCACACACATCACACCTTACCATTAATGAGCTCC 13042

QY 145 GACTTCATTTACACCTCGATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 204

DB 13041 ACACATCAATATCTCTATATGAGTCCACGACGACGCTTTCACATCACCACACGACC 12982

QY 205 ACTCTGCTCTCTGCGACCATCTAATGATGCTCCGAAAGCGACGCTCCCATCCCATTC 264

DB 12981 ACAACCAACACCTATCATCTGACTTCACCATCAATCAATTCATCATCTCCACGAC 12922

QY 265 GATTTCTACCA 275

DB 12921 CATCATCACA 12911

RESULT 3

US-09-949-016-12843/c

Sequence 12843, Application US/09949016

Patent No. 6812339

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

CURRENT APPLICATION NUMBER: US/09/949,016

PRIOR FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 12843

LENGTH: 57280

TYPE: DNA

ORGANISM: Human

US-09-949-016-12843

Query Match 2.1%; Score 50.2; DB 3; Length 57280;

Best Local Similarity 53.9%; Pred. No. 0.0059;

Matches 103; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 85 AGCCACAACACCTCTACACATCTGCTCCGCGACGAAATGGCGGACCGCTCTCTCTCC 144

DB 13101 AGCACCATACACACACCAAGCCCTCCACACACATCACACCTTACCATTAATGAGCTCC 13042

QY 145 GACTTCATTTACACCTCGATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 204

DB 13041 ACACATCAATATCTCTATATGAGTCCACGACGACGCTTTCACATCACCACACGACC 12982

QY 205 ACTCTGCTCTCTGCGACCATCTAATGATGCTCCGAAAGCGACGCTCCCATCCCATTC 264

DB 12981 ACAACCAACACCTATCATCTGACTTCACCATCAATCAATTCATCATCTCCACGAC 12922

QY 265 GATTTCTACCA 275

DB 12921 CATCATCACA 12911

RESULT 4

US-09-949-016-12844/c

Sequence 12844, Application US/09949016

Patent No. 6812339

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

CURRENT APPLICATION NUMBER: US/09/949,016

PRIOR FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 12844

LENGTH: 57280

TYPE: DNA

ORGANISM: Human

US-09-949-016-12844

	Matches	103;	Conservative	0;	Mismatches	88;	Indels	0;	Gaps	0;
QY	85	AGCCCAACACCTCTACACTATCTGTCGCGACGAAATGGCCGACCGCTTCTCTCC								144
Db	13101	AGCACCAATACACACCAACATGGCTCTCCACACCATACACCACTTACCACTTATGAGCTTC								13042
QY	145	GACTTCATTTTCACTCCGATTCCTCTCTCTCTCTTGGCCACGCGACCAACCGCC								204
Db	13041	ACCATCAATATCCCTATCAAGTTCACACCAACCGCTTTCACATACACACCAACACACC								12982
QY	205	ACTCTGCTCTCTGTGCACCATATATGATGTCGCCAAGGCCACGTGCCATATCCCAT								264
Db	12981	ACAACGACCACTCACTATCATGACTTCCACACATCAACATTAACATCACTCCACACAAC								12922
QY	265	GATTTCTACCA	275							
Db	12921	CATCATCACA	12911							

```

RESULT 8
US-09-949-016-13544/c
; Sequence 13544, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 13544
; LENGTH: 57280
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13544

Query Match          2.1%; Score 50.2; DB 3; Length 57280;
Best Local Similarity 53.9%; Pred. No. 0.0059;
Matches 103; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY      85  AGCCACACACCTCTACAACTATCTGCTCCGCCAGCAAAATGGCCGACCGCTCTTCTCTC 144
        |||||
Db       13101 AGCACCATCACACCCACCAACATCGCTCCACACCATCACACCTTCAACATTAAGACTCC 13042

QY      145  GACCTCAATTTTCAACCTCGATTCCTCTCTCTCTTGCCACGCGCACACACGCGC 204
        |||||
Db       13041 ACCATCAATATCTCTATACGCTCACACACACCGTTTACACATCACACACACACAC 12982

QY      205  ACTCTGCTCTCTCTGCGCAACCATCTATTTGAATGCTCCGAAAGCCACGCTCCCATCCCATTT 264
        |||||
Db       12981 ACAACACACACCTCACATCATGACTTCCACGATCAACATTACATCAACCTCCACCAAC 12922

QY      265  GATTTCATCCA 275
        |||||
Db       12921 CATCATCACGA 12911

RESULT 9
US-09-949-016-13545/c
; Sequence 13545, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

```

```

FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949, 016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ. ID NOS: 207012
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ. ID NO 13545
LENGTH: 57280
TYPE: DNA
ORGANISM: Human
US-09-949-016-13545

Query Match      2.1%; Score 50.2; DB 3; Length 57280;
Best Local Similarity 53.9%; Pred. No. 0.0059;
Matches 109; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY      85 AGCCACAAACCTCTTCAACTATATCTGCTCCGCGACGAATGGAGCGACCGCTTCTTCTTC 144
Db      13101 AGACCAATTCACCAACAACTGGCTTCACACACCATATCACACCTTACCACTTATGAGCTTC 130422

QY      145 GACTTCATTTGCACTTCGCCATTCCTCTCTCTCTCTTGGCAGCGCCACCAACCGCC 204
Db      13041 ACCATCATATTCCTCTTACAGTTCACACCAACCGCTTCTTACCATATACACCAACCAACAC 129882

QY      205 ACTCTGCTCTCTGCGACCACTATTTGATCGTCCGAAAGCCACAGTCCCATTCGCCAATT 264
Db      12981 ACAACCAACCACTCATATCATGACTTCACACCACTAACATTTACCATCATCCTCCACCAAC 129222

QY      265 GATTTCTACCA 275
Db      12921 CATCATTCGCA 12911

```

```

RESULT 10
US-09-949-016-14633/C
; Sequence 14633, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C0001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14633
; LENGTH: 57280
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14633

Query Match      2.1%; Score 50.2; DB 3; Length 57280;
Best Local Similarity 53.9%; Pred. No. 0.0059;
Matches 103; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

OY      85  AGCCACACACCTCTACACATCATGCTCCGCCAGACAATGAGCGCTCTCTCC 144
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      13101 AGCACATCACACACACCACTAGCTCTCCACACACATCACACACTTACCATTTATGAGTCC 13042

OY      145  GACTTCAATTTCACCTCCGATTCTCTCTCTCTCTCTTTCGCGACCGCCACACACACCGCC 204
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Db	13041	ACCATCATATATCCCTTATCAAGTCACACACACACCGGTCTTACCATCAACACACACACACAC	12982
Qy	205	ACTCTGCTCTCTCTGCGACCATATATGATATGTCGCCAAGACAGTCGCCATATCCCATTT	264
Db	12981	ACAAACACACCACTCAGCATATCATGACTTCCACCATCAACATTAACATCACTCCACCCACAC	12922
Qy	265	GATTTCTACCA	275
Db	12921	CATCATCAGCA	12911

```

RESULT 11
US-09-949-016-14634/c
; Sequence 14634, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14634
; LENGTH: 57280
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-14634

```

```

RESULT 12
US-09-949-016-14635/C
; Sequence 14635, Application US/09949016
; Patent No. 681239
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498

```

```

; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ. ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14635
; LENGTH: 57280
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14635

```

```

RESULT 13
US-09-949-016-14636/c
; Sequence 14636, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C0001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14636
; LENGTH: 57280
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14636

```

	Query Match	Score	DB 3	Length
	2.1%	50.2	57280	
	Best Local Similarity	53.9%	Prod. No. 0.0059	
	Matches	109	Conservative	0
			Mismatches	88
			Indels	0
			Gaps	0
Qy	85	AGCCACAAACCTCTTCAACTATCTGCTCCGACGAAATGGAGCCGCGTCTTCTTCC	144	
Db	13101	AGACACATTCACACCAACACTGCTTCCACACCATCACACCTTCAACATTTAGAGCTCC	13042	
Qy	145	GACCTCAATTTCACTCCGATTCCTCTCTCTCTCTTGGCAGCGCCACGACCGCC	204	
Db	13041	ACCATCAATATTCCTTATCAAGTCCACACACCGTCTTATCATTACATCACACGACACCC	12982	
Qy	205	ACCTCGTCTCTGTGGCACCATTTATGATGCTGCGGAAGCGACAGTCCCATCCCATTT	264	
Db	12981	ACGACACACGACTACTATGATGACTTCCACCATTAACATTCATCCTCCACACCAAC	12922	
Qy	265	GATTTCTACCA	275	

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 10, 2005, 16:57:24 ; Search time 11911.1 Seconds
(without alignments)
11482.131 Million cell updates/sec

Title: US-10-600-070B-1

Perfect score: 2406
Sequence: 1 atggaagcttcgagtcacgt.....gtctgtcttcgacataa 2406

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb_da:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pac:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_sts:*
11: gb_sy:*
12: gb_un:*
13: gb_vl:*
14: gb_htg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2402.8	99.9	2438	15	AY221469 Arabidops
2	2399.6	99.7	2436	15	AY150490 Arabidops
3	2399.6	99.7	2637	15	AY091075 Arabidops
4	1724	71.7	3668	15	AY221468 Arabidops
5	1720.8	71.5	3664	15	AY221467 Arabidops
6	1717.6	71.4	85791	15	AB016888 Arabidops
7	243.6	10.1	133779	14	AC158210 Medicago
8	180.8	7.5	110000	15	AP008208_011
9	180.8	7.5	150462	15	AP004885 Oryza sat
10	165.4	6.9	550	10	DQ022571 Beta vulg
11	163.4	6.8	184752	14	AC155597 Zea mays
12	144.2	6.0	117818	14	AC160013 Medicago
13	62	2.6	221756	14	AC094893 Rattus no
14	62	2.6	232974	14	AC114193 Rattus no
15	60.8	2.5	200412	14	AC115666 Rattus no
16	60.8	2.5	253094	14	AC113774 Rattus no
17	60.2	2.5	23373	14	AC134805 Rattus no
18	59.6	2.5	193341	14	AC134500 Rattus no

19	59.2	2.5	96573	14	AC136672 Rattus no
20	59.2	2.5	175351	14	AC117900 Rattus no
21	59.2	2.5	233591	14	AC121387 Rattus no
22	59.2	2.5	243553	14	AC096320 Rattus no
23	59.2	2.5	247341	14	AC099178 Rattus no
24	59.2	2.5	312143	14	AC123509 Rattus no
25	59.2	2.5	346208	14	AC128290 Rattus no
26	57.6	2.4	7218	6	166494 Sequence 14
27	57.6	2.4	255678	14	AC099432 Rattus no
28	57.6	2.4	313096	14	AC11879 Rattus no
29	57.2	2.4	98521	9	AC094507 Rattus no
30	57	2.4	263036	14	AC105597 Rattus no
31	56.8	2.4	110000	14	CR954204_0 Oestreococ
32	56.4	2.3	282355	14	AC104401 Rattus no
33	56	2.3	143976	14	AC114543 Rattus no
34	56	2.3	246199	14	AC119558 Rattus no
35	56	2.3	252996	14	AC130153 Rattus no
36	56	2.3	253772	14	AC109777 Rattus no
37	55.6	2.3	242516	14	AC106523 Rattus no
38	54.6	2.3	165849	14	AC114251 Rattus no
39	54.6	2.3	231716	9	AC135409 Rattus no
40	54.6	2.3	234812	14	AC111482 Rattus no
41	54	2.2	110000	14	AC129796_0 Rattus no
42	54	2.2	240800	14	AC095964 Rattus no
43	54	2.2	274697	14	AC098240 Rattus no
44	53.8	2.2	205796	14	AL713891 Mus muscu
45	53.8	2.2	215524	9	AL603913 Mouse DNA

ALIGNMENTS

RESULT 1
AY221469 2438 bp mRNA linear PLN 05-AUG-2003
LOCUS Arabidopsis thaliana division protein (ARCF) mRNA, complete cds;
DEFINITION nuclear gene for chloroplast product.
ACCESSION AY221469
VERSION AY221469.1 GI:33436353
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE 1 (bases 1 to 2438)
Vitha,S., Froehlich,J.E., Koksharova,O., Pyke,K.A., Van Erp,H. and Oosteryoung,K.W.
ARC6 is a J-Domain Plastid Division Protein and an Evolutionary Descendant of the Cyanobacterial Cell Division Protein Ftn2
JOURNAL Plant Cell 15 (8), 1918-1933 (2003)
PUBMED 12897262
REFERENCE 2 (bases 1 to 2438)
Vitha,S., Koksharova,O., van Erp,H., Froehlich,J.E. and Oosteryoung,K.W.
Direct Submission
Submitted (17-JAN-2003) Department of Plant Biology, Michigan State University, 166 Plant Biology Building, East Lansing, MI 48824, USA
FEATURES
source
location/Qualifiers
1. 2438
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/ecotype="Wassilewskija"
1. 2438
/gene="ARC6"
/note="synonym: At5g42480"
2. 2407
/gene="ARC6"
/codon_start=1
/product="division protein"
/protein_id="AA018646.1"
/db_xref="GI:33436354"

```
/translation="MEALSHVIGLSPFOCLRLEPATYTKLRSHNTSTTICSAKMD
RLISDFNPTSSSSSPATATATATATLSPSIDRPERHPIDPFQVIGAOHPFT
DJIRAPEARVSKPPOGSDALISROLOAECTLSRSPRSRREGLDDEBGT
VITDVPWDKVGALCTOEGEETIVRVEGALKERLPSPDQVILWALAFVLDG
RDMAALDPPDPTCYFEVEERKILQSEGNASSLAPDLRAQIDLEBETITRYVLELG
LPUDDDYAKRNLKLSGVRNLSVGGGASALVGLJREFPNKATLRAVLENSND
FVATPSPNIPASSFEVYVALALVAQAFIGKPHILOADQFOLOQAKWAMEIPM
LYDRNNMEIDFGLERGLCALIGKVECKWMLGLDSEDSQYRNPAIVEVLENSND
DNDLPGLCKLEETWLAGVVEPRPRDKKFKLGDYDDBMYLSYERVEVQSP
AAATMARIGAEHYKASAMOLKVPSPRYTDRNSAPKPOVETFEVSDVGNVGD
GEBGVPIABAVRPRENFTNDYAIRAVSSSDVETVEVSVAMLEKASVKILAAV
AIGLISLPSQIKLSSSSFOKQDWSMESDVATIGSVRADSEALPRMDIAENI
VSKWQIKSLAFGPDRIEMLPVILDRMKIWTDRAEATQLOLVYDYLKLSVDS
VTVASDSTRALVEATEESACLSLVHPENNATDVRTYTRIEVFMKSGMKITGSGV
LAS"
```

misc_difference 474
/gene="ARC6"
/note="Results in Gly to Glu substitution"
/replace="a"
misc_difference 1831
/gene="ARC6"
/note="Results in Glu to Asp substitution"
/replace="t"

ORIGIN

Query Match 99.9%; Score 2402.8; DB 15; Length 2438;
Best Local Similarity 99.9%; Pred. No. 0;

Matches 2404; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 ATGGAACCTGAGTCACTGCGCATTTGGTCTCTCCCATTCCAATTATGCCATTAACA 60
DB 2 ATGGAACCTGAGTCACTGCGCATTTGGTCTCTCCCATTCCAATTATGCCATTAACA 61
QY 61 CCGGCGACGACAAAGCTCCGACGTAGCCACAACACCTCTACACATATCTGCTCCGACG 120
DB 62 CCGGCGACGACAAAGCTCCGACGTAGCCACAACACCTCTACACATATCTGCTCCGACG 121
QY 121 AATGGGCGGACCGTCTTCTCTCGACTTCAATTTACCTCCGATCTCTCTCTCTCC 180
DB 122 AATGGGCGGACCGTCTTCTCTCGACTTCAATTTACCTCCGATCTCTCTCTCTCC 181
QY 181 TTGGCCACCGGACACACACCGGACACTCTGCTCTCTGCGACCATTAATGATGCTCC 240
DB 182 TTGGCCACCGGACACACACCGGACACTCTGCTCTCTGCGACCATTAATGATGCTCC 241
QY 241 GAAGCGACGCGCCCATCCCATGATTTCTACACAGGATTAAGGACCAAAACATTTTC 300
DB 242 GAAGCGACGCGCCCATCCCATGATTTCTACACAGGATTAAGGACCAAAACATTTTC 301
QY 301 TTAAACGATGGAATCAGAAAGCAATCGAAGCTAGGGTTTCGAAACCGCGCAATTCGCT 360
DB 302 TTAAACGATGGAATCAGAAAGCAATCGAAGCTAGGGTTTCGAAACCGCGCAATTCGCT 361
QY 361 TTACGCGACGACGCTTTATATACGCCGAGACAGATTTCTTCAAGCTGCTTGGCAAACTCTG 420
DB 362 TTACGCGACGACGCTTTATATACGCCGAGACAGATTTCTTCAAGCTGCTTGGCAAACTCTG 421
QY 421 TCTAATCTCGGTCTAGAAAGAGATACATGAAGAGTCTTCTTGATGATGAAGAGCTACA 480
DB 422 TCTAATCTCGGTCTAGAAAGAGATACATGAAGAGTCTTCTTGATGATGAAGAGCTACA 481
QY 481 GTCATCATGATGTTCTTGGGATTAAGGTTCTTGAGGCTCTCTGTGTATTTGCAAGAGGT 540
DB 482 GTCATCATGATGTTCTTGGGATTAAGGTTCTTGAGGCTCTCTGTGTATTTGCAAGAGGT 541
QY 541 GGTGAGACTGAGATAGTTCTTCCGCTTGTGAGGCTCTGCTTTAAGAGAGGTTCCCTAAG 600
DB 542 GGTGAGACTGAGATAGTTCTTCCGCTTGTGAGGCTCTGCTTTAAGAGAGGTTCCCTAAG 601
QY 601 TCGTTTAAGCAAGATGCTTTTATGATTATGGCGCTTGCTTCTGATGCTCGAGGAGAT 660
DB 602 TCGTTTAAGCAAGATGCTTTTATGATTATGGCGCTTGCTTCTGATGCTCGAGGAGAT 661
```

```
QY 661 GCTATGCGATTGATCCACCTGATTTTATTAATCTGTTATAGTTTGTGAGAGCTTTG 720
DB 662 GCTATGCGATTGATCCACCTGATTTTATTAATCTGTTATAGTTTGTGAGAGCTTTG 721
QY 721 AAGCTTTTACAGAGAGAGAGCAAGTAGCTTGCACCGGATTTTACGTGACAAATTTGAT 780
DB 722 AAGCTTTTACAGAGAGAGAGCAAGTAGCTTGCACCGGATTTTACGTGACAAATTTGAT 781
QY 781 GAGACTTTGGAAGAGATCACTCCGCGTTATGCTTGTGAGCTACCTTGAGCTTGTG 840
DB 782 GAGACTTTGGAAGAGATCACTCCGCGTTATGCTTGTGAGCTACCTTGAGCTTGTG 841
QY 841 GATGATTAACGCTCGAAGAAAGCTAAATGCTTTAAGCGGTGCGGAATATTTTGCTCT 900
DB 842 GATGATTAACGCTCGAAGAAAGCTAAATGCTTTAAGCGGTGCGGAATATTTTGCTCT 901
QY 901 GTTGAAGAGAGTGAAGATCACTCTTGTGTTGGGGTTTGAACCCGTGAGAAATTATGAAT 960
DB 902 GTTGAAGAGAGTGAAGATCACTCTTGTGTTGGGGTTTGAACCCGTGAGAAATTATGAAT 961
QY 961 GAGGCGTTTTTACGAATGACAGCTGCGAGCAGGTTGATCTTTTGTAGCTACCCCAAG 1020
DB 962 GAGGCGTTTTTACGAATGACAGCTGCGAGCAGGTTGATCTTTTGTAGCTACCCCAAG 1021
QY 1021 AATATTCACAGACAGTCAATTTGAAGTTTACGAAGTTGCACTTGTCTTGTGCTCAAGCT 1080
DB 1022 AATATTCACAGACAGTCAATTTGAAGTTTACGAAGTTGCACTTGTCTTGTGCTCAAGCT 1081
QY 1081 TTATATGTGAAGAGCACACCTTTTACAGAGCTGATPAAGAAATTCAGCAACTTCAG 1140
DB 1082 TTATATGTGAAGAGCACACCTTTTACAGAGCTGATPAAGAAATTCAGCAACTTCAG 1141
QY 1141 CAGGCTAAGGTAATGCTATGAGATTCCTGCGATGTTGATGATACACGAAATTAATGG 1200
DB 1142 CAGGCTAAGGTAATGCTATGAGATTCCTGCGATGTTGATGATACACGAAATTAATGG 1201
QY 1201 GAGATAGACTTCGCTCTAGAAAGGAGACTCTGTGCACTGCTTATAGGCAAAAGTTGATGA 1260
DB 1202 GAGATAGACTTCGCTCTAGAAAGGAGACTCTGTGCACTGCTTATAGGCAAAAGTTGATGA 1261
QY 1261 TGGCGTATGTTGGGCTTGAACATGAGAGATTCACAAATATGGAATCCAGTATTTGG 1320
DB 1262 TGGCGTATGTTGGGCTTGAACATGAGAGATTCACAAATATGGAATCCAGTATTTGG 1321
QY 1321 GAGTTGTTTGGAGAAATCAATTCGTATGACAAATGATATCTCCGTGACATATGCAAA 1380
DB 1322 GAGTTGTTTGGAGAAATCAATTCGTATGACAAATGATATCTCCGTGACATATGCAAA 1381
QY 1381 TTGTTGGAACCTGTTGGCAGGGGTTGCTTTCTTCTAGGTTTCAGAGACACCAAGATTA 1440
DB 1382 TTGTTGGAACCTGTTGGCAGGGGTTGCTTTCTTCTAGGTTTCAGAGACACCAAGATTA 1441
QY 1441 AATTTTAACTCGGGGACATATATGATATCCATAGGTTTGAAGTTTGAAGAAAGTG 1500
DB 1442 AATTTTAACTCGGGGACATATATGATATCCATAGGTTTGAAGTTTGAAGAAAGTG 1501
QY 1501 GAGGTAGTTCAAGGTTCTCTTTAGCTGCTGCAACTATGCGAAGATTGGAGCGAG 1560
DB 1502 GAGGTAGTTCAAGGTTCTCTTTAGCTGCTGCAACTATGCGAAGATTGGAGCGAG 1561
QY 1561 CATGTAAAGCTATGCTATGACGAGCACTGACAGAAAGTTTCTTCCGCTATATACAT 1620
DB 1562 CATGTAAAGCTATGCTATGACGAGCACTGACAGAAAGTTTCTTCCGCTATATACAT 1621
QY 1621 AGAACTCGGCTGAAACCAAGATGACAGAGACAGTGTATAGTATGATCCGTGGT 1680
DB 1622 AGAACTCGGCTGAAACCAAGATGACAGAGACAGTGTATAGTATGATCCGTGGT 1681
QY 1681 AACAAATGAGCCGCTATGAGTGAAGCTGCTGCTTATATTTGCAAGACTGTAAACCTCT 1740
DB 1682 AACAAATGAGCCGCTATGAGTGAAGCTGCTGCTTATATTTGCAAGACTGTAAACCTCT 1741
QY 1741 GAAAACTTTGAAACTATATGATTAATGCAATTCGAGTGGGGCTTCAGAGAGTACGTTGAT 1800
```

```

Db      1742 GAAACCTTGGAACTAATGATATCAATTCAGCTGGGCTTCAGAGATGACGTTGAT 1801
Qy      1801 GAACTACTGTTGAAATGTCCTGTCGTATATGTTAAAGAGCAGTGTGAAGATCCTA 1860
Db      1802 GAACTACTGTTGAAATGTCCTGTCGTATATGTTAAAGAGCAGTGTGAAGATCCTA 1861
Qy      1861 GCTGCTGCTGGCAATGGAATGATTCACCTGTCACCCAGAGATATTTCTTAAAGC 1920
Db      1862 GCTGCTGCTGGCAATGGAATGATTCACCTGTCACCCAGAGATATTTCTTAAAGC 1921
Qy      1921 AGCTCATCTTTTCAACGCAAGATATGTTTCTTAAAGATCTGAATGCTGCTACCA 1980
Db      1922 AGCTCATCTTTTCAACGCAAGATATGTTTCTTAAAGATCTGAATGCTGCTACCA 1981
Qy      1981 GGGTCAGTCAGAGCTGAGATTCAGAGACATCTCCAGAGATGATGATAGGCTGCAAG 2040
Db      1982 GGGTCAGTCAGAGCTGAGATTCAGAGACATCTCCAGAGATGATGATAGGCTGCAAG 2041
Qy      2041 AATATAGTATCCAACTGGCAGAGATTAAGTCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
Db      2042 AATATAGTATCCAACTGGCAGAGATTAAGTCTGCTGCTGCTGCTGCTGCTGCTGCT 2101
Qy      2101 GAAATGTTACCAAGCTTTTGGATGGCGAATGCTGAAGATTTGGACTGACAGACAGCT 2160
Db      2102 GAAATGTTACCAAGCTTTTGGATGGCGAATGCTGAAGATTTGGACTGACAGACAGCT 2161
Qy      2161 GAACTGCGCAGCTGCTGGCTTATGATTAATACACTGTTGAAACCTATCTGTTGACAGT 2220
Db      2162 GAACTGCGCAGCTGCTGGCTTATGATTAATACACTGTTGAAACCTATCTGTTGACAGT 2221
Qy      2221 GTGACAGTCTCAGCAGATGAGAACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2280
Db      2222 GTGACAGTCTCAGCAGATGAGAACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2281
Qy      2281 TGTCTATCTGATTTTGGCTTATCCAGAAAACAATGCTACTGATGTCAGAACTTACACACA 2340
Db      2282 TGTCTATCTGATTTTGGCTTATCCAGAAAACAATGCTACTGATGTCAGAACTTACACACA 2341
Qy      2341 AGATACGAAAGTTTCTGGTCCAGCTCAGGGTGAAGAAATCACTGAGAGCTGTTCTTGA 2400
Db      2342 AGATACGAAAGTTTCTGGTCCAGCTCAGGGTGAAGAAATCACTGAGAGCTGTTCTTGA 2401
Qy      2401 TCATTA 2406
Db      2402 TCATTA 2407

```

RESULT 2
AY150490
LOCUS Arabidopsis thaliana unknown protein (At5g42480) mRNA, complete cds.
DEFINITION Arabidopsis thaliana unknown protein (At5g42480) mRNA, complete cds.
ACCESSION AY150490
VERSION AY150490.1 GI:33297711
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
REFERENCE 1 (bases 1 to 2436)
AUTHORS Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M., Quach,H.L., Tang,C., Toriumi,M., Wallender,E.K., Wong,C., Wu,H.C., Yu,G., Yuan,S., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Kawai,T., Kim,C.J., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shim,P., Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W., Becker,J.R. and Theologis,A.
TITLE Arabidopsis Open Reading Frame (ORF) Clones
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2436)
AUTHORS Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M.,

Quach,H.L., Tang,C.C., Toriumi,M., Wallender,E.K., Wong,C., Wu,H.C., Yu,G., Yuan,S., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Kawai,T., Kim,C.J., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shim,P., Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W., Becker,J.R. and Theologis,A.
 Direct Submission
 Submitted (13-SEP-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA
 The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-length CDNA') : Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGEC (SSP) Consortium members constructed and sequenced the pUNI (ORF) clones using the RAFL cDNAs: Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M., Quach,H.L., Tang,C.C., Toriumi,M., Wallender,E.K., Wong,C., Wu,H.C., Yu,G., Yuan,S., Carninci,P., Chen,H., Cheuk,R., Jones,T., Kim,C.J., Nguyen,M., Palm,C.J., Shim,P., Southwick,A., Tripp,M.G., Wu,T., Davis,R.W., Becker,J.R. and Theologis,A.

Yamada,K. (SSP/PGEC) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP/PGEC) contributed equally to this work as PI's.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

FEATURES
 Location/Qualifiers
 1..2436
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /chromosome="5"
 /clone="U19395"
 /ecotype="Columbia"
 /note="This clone is in pUNI 51."
 1..2436
 /gene="At5g42480"
 1..2406
 /gene="At5g42480"
 /codon_start=1
 /evidence=experimental
 /product="unknown protein"
 /protein_id="AA12907.1"
 /db_xref="GI:23297712"
 /translation="MEALSHVIGTGLSPFQCLRPATTKLRSHNTSTTCSAKWADRLSDFNFTSDSSSPATATTTATLVSPPSIDRPRRHVPIPIDFYQVIGAQTHPLT DGJRAFEARVSRPQGFSDDALISRQILOAACEITSPRSREYNEGILADEEAT VITDVPMDKVGALCVLOEGGETEIVLRGVEALLKERTLPSFKQDVVLAALAFDVS RDMALDLPDPTGTFEYFEALATKLDEGASSLAPDRADIDETLETETPTPVYUEILG LPLGDTAAKRLNGLSGRTIILSVGGGASALVGCITREKFMARERLRTAAEVDVL FVATPSNIPAESFEYEVVALALVAQAFIGKPHLLODADKQFOLOQAQWAMEIPM LLYDTRNNWEIDFGLERGLCALIGKDECEMMLGLDESDQYRPAIVEFLVLSNMD DNDLDELCLLETLMLAGVVEPRFDKDKFLGDYDPMVLSYLERVAVVSGPL AAAAAMARICAEHVKASAMQALQKVPESRYTRDNRSAEPKQVOTFVSVDVGNVGD GECQVETAEARVPSNPETMDVATRAQVSSVDETVMSVADMEKASVKTAAAY ATGILISLFSQKTYLTKSSSFQRKQMSMESDVATIGSVRADDSALPRMDARTAEI VSWQKIKSLAFPDHRIEMLPVLIDGRMLKITDRAEATQAGLVVDYTLTKLSVDS VTYWSAGTQRLVATLEESACLSDLVHPENNATDVRYTTRTREVFSKSKWKITTEGVS LAG"
 2407..2436
 /gene="At5g42480"
 misc_difference 2433..2434
 /gene="At5g42480"
 /note="compared to genomic sequence"
 /replace="ttg"

ORIGIN
 Query Match 99.7%; Score 2399.6; DB 15; Length 2436;
 Best Local Similarity 99.8%; P-Val. No. 0;
 Matches 2402; Conservative 4; Indels 0; Gaps 0;

QY 1 ATGGAAGCTGAGTCAAGTCGAGTATGATCTCTCCCATTCGAATTATGCCATTCACA 60
Db 1 ATGGAAGCTGAGTCAAGTCGAGTATGATCTCTCTCCCATTCGAATTATGCCATTCACA 60
QY 61 CCGGCGAAGCAAAAGCTCCGAGTGAAGCAACAAGCTCTTACAACTATCTGCTCCGACG 120
Db 61 CCGGCGAAGCAAAAGCTCCGAGTGAAGCAACAAGCTCTTACAACTATCTGCTCCGACG 120
QY 121 AAATGGGCGGACGGCTTCTCTCGGACTTCAATTCACTCCGATTCCTCTCTCTCC 180
Db 121 AAATGGGCGGACGGCTTCTCTCGGACTTCAATTCACTCCGATTCCTCTCTCTCC 180
QY 181 TTGCGCAACCGCAACCAACCGCAACTCTGCTCTCTGCGCAACCATCTATTTGATGTC 240
Db 181 TTGCGCAACCGCAACCAACCGCAACTCTGCTCTCTGCGCAACCATCTATTTGATGTC 240
QY 241 GAAGCGCAAGTCCCATTCCTCCATGATTTCTACAGGATTTAGAGCTCAAAACATTTTC 300
Db 241 GAAGCGCAAGTCCCATTCCTCCATGATTTCTACAGGATTTAGAGCTCAAAACATTTTC 300
QY 301 TTAAACCAATGGAATCAAGAGCAATGGAAGTGAAGGTTTCGAAACCGCGCAATTCG 360
Db 301 TTAAACCAATGGAATCAAGAGCAATGGAAGTGAAGGTTTCGAAACCGCGCAATTCG 360
QY 361 TTGACGACAGACGCTTTATATAGCCGAGACAGATTTCTCAAGCTGCTTGGAACCTG 420
Db 361 TTGACGACAGACGCTTTATATAGCCGAGACAGATTTCTCAAGCTGCTTGGAACCTG 420
QY 421 TCTAATCTCGGTCTAGAGAGAGTATCAATGAAGTCTTCTTGATGATGAAGAGTACA 480
Db 421 TCTAATCTCGGTCTAGAGAGAGTATCAATGAAGTCTTCTTGATGATGAAGAGTACA 480
QY 481 GTCAATCAATGATCTCTGAGTGAAGTTCCTGAGGCTCTCTGATATTCGAAGAGT 540
Db 481 GTCAATCAATGATCTCTGAGTGAAGTTCCTGAGGCTCTCTGATATTCGAAGAGT 540
QY 541 GGTGAGA CTGAGATAGTCTTCGAGGTGCTGAGGCTCTGCTTMAAGAGAGTTCCTAG 600
Db 541 GGTGAGA CTGAGATAGTCTTCGAGGTGCTGAGGCTCTGCTTMAAGAGAGTTCCTAG 600
QY 601 TCGTTAAGCAAGATGCTGTTTATGATGAGGCTTCGCTTTCGATGCTCGAGGAT 660
Db 601 TCGTTAAGCAAGATGCTGTTTATGATGAGGCTTCGCTTTCGATGCTCGAGGAT 660
QY 661 GCTATGAGATGATTCACCTGATTTTATCTGATATGAGTTCGTTGAGGAGCTTTCG 720
Db 661 GCTATGAGATGATTCACCTGATTTTATCTGATATGAGTTCGTTGAGGAGCTTTCG 720
QY 721 AAGCTTTTACAGAGAGAGCAAGTATGCTTCACCGAGTTTACGTCACAAATTTGAT 780
Db 721 AAGCTTTTACAGAGAGAGCAAGTATGCTTCACCGAGTTTACGTCACAAATTTGAT 780
QY 781 GAGACTTTGAGAGAGATCACTCCGCGTATGCTTTGAGCTACTTGGCTTACCGCTTGT 840
Db 781 GAGACTTTGAGAGAGATCACTCCGCGTATGCTTTGAGCTACTTGGCTTACCGCTTGT 840
QY 841 GATATTTACGCTGGAAGAGCACTAAATGCTTAAAGCGGTGCGGAATTTTGTGCT 900
Db 841 GATATTTACGCTGGAAGAGCACTAAATGCTTAAAGCGGTGCGGAATTTTGTGCT 900
QY 901 GTTGGAGAGGTGAGCATCTCTGTTGGGGTTTGAACCGTGAAGATTTATGAT 960
Db 901 GTTGGAGAGGTGAGCATCTCTGTTGGGGTTTGAACCGTGAAGATTTATGAT 960
QY 961 GAGGCGTTTTTACAGATGACAGTCTGAGAGAGTTCCTTTTGTAGTACCCCAAGC 1020
Db 961 GAGGCGTTTTTACAGATGACAGTCTGAGAGAGTTCCTTTTGTAGTACCCCAAGC 1020
QY 1021 AATTTTCCAGAGAGTCAATTTGAGTTTACAGAGTTCCTTTGAGTCAAGCT 1080
Db 1021 AATTTTCCAGAGAGTCAATTTGAGTTTACAGAGTTCCTTTGAGTCAAGCT 1080

QY 1081 TTTATTTGTAAGAGACCAACCTTTTACAGATGCTGATTAAGCAATTCAGCAACTTCAG 1140
Db 1081 TTTATTTGTAAGAGACCAACCTTTTACAGATGCTGATTAAGCAATTCAGCAACTTCAG 1140
QY 1141 CAGGCTTAAGGTAATGCTATGAGATTTCTGCGAATTTGTATGATACAGGAATTAATTCG 1200
Db 1141 CAGGCTTAAGGTAATGCTATGAGATTTCTGCGAATTTGTATGATACAGGAATTAATTCG 1200
QY 1201 GAGATAGACTTCGGTCTAGAGAGAGGAGCTCTGTCATCTGCTTATAGGCAAGTTGATGA 1260
Db 1201 GAGATAGACTTCGGTCTAGAGAGAGGAGCTCTGTCATCTGCTTATAGGCAAGTTGATGA 1260
QY 1261 TGCCTATGATGTTGGGCTTTAGACAGTGAAGATTCACAAATATAGCAATTCAGTATTCG 1320
Db 1261 TGCCTATGATGTTGGGCTTTAGACAGTGAAGATTCACAAATATAGCAATTCAGTATTCG 1320
QY 1321 GAGTTGTTTGGAGAAATTCAAATGCTGATGACAAATGATGATTCCTGCTGACTATGCAAA 1380
Db 1321 GAGTTGTTTGGAGAAATTCAAATGCTGATGACAAATGATGATTCCTGCTGACTATGCAAA 1380
QY 1381 TTGTTGAAAACTGCTGTTGGCAGGGGTTGCTTTCCTAGGTTCAAGACACAAAGATTA 1440
Db 1381 TTGTTGAAAACTGCTGTTGGCAGGGGTTGCTTTCCTAGGTTCAAGACACAAAGATTA 1440
QY 1441 AAATTTAACTCGGGAAGTATATGATGATCTTATGATGATCTTGAAGTTCCTTGAAGAGTG 1500
Db 1441 AAATTTAACTCGGGAAGTATATGATGATCTTATGATGATCTTGAAGTTCCTTGAAGAGTG 1500
QY 1501 GAGGTAGTTCAAGGTTCTCTTTAGCTGCTGTCGCAATATGCAAGATTTGAGAGCCGAG 1560
Db 1501 GAGGTAGTTCAAGGTTCTCTTTAGCTGCTGTCGCAATATGCAAGATTTGAGAGCCGAG 1560
QY 1561 CATGTGAAGCTAGTGTATGACAGGACATGACAGAAAGTTTTCCTCCGCTATACAGAT 1620
Db 1561 CATGTGAAGCTAGTGTATGACAGGACATGACAGAAAGTTTTCCTCCGCTATACAGAT 1620
QY 1621 AGAACTCGGCTGACCCCAAGAGATGCAAGACAGTGTATAGTATGATCTGTTGGT 1680
Db 1621 AGAACTCGGCTGACCCCAAGAGATGCAAGACAGTGTATAGTATGATCTGTTGGT 1680
QY 1681 AACAATGATGAGCCGTATGAGTGAAGCTGCTGATCTTATGAGAGAACTGTAAAGCCCTCT 1740
Db 1681 AACAATGATGAGCCGTATGAGTGAAGCTGCTGATCTTATGAGAGAACTGTAAAGCCCTCT 1740
QY 1741 GAAAACCTTGAAGAACTATGATTTATGCAATTCGAGTGGGCTTCAGAGAGTACGTTGAT 1800
Db 1741 GAAAACCTTGAAGAACTATGATTTATGCAATTCGAGTGGGCTTCAGAGAGTACGTTGAT 1800
QY 1801 GAAACTATCTGTTGAAATGTCCTGCTGATATGTTAAAGAGGCAAGTGTGAAGATCTTA 1860
Db 1801 GAAACTATCTGTTGAAATGTCCTGCTGATATGTTAAAGAGGCAAGTGTGAAGATCTTA 1860
QY 1861 GCTGCTGCTGAGCAATGAGCTGATTTCACTGTTAGGCGAGAGATTTTCTTAAGC 1920
Db 1861 GCTGCTGCTGAGCAATGAGCTGATTTCACTGTTAGGCGAGAGATTTTCTTAAGC 1920
QY 1921 AGCTCATCTTTTCAAGCAAGAGATATGTTTCTTATGAGATCTGATGCTGCTACATA 1980
Db 1921 AGCTCATCTTTTCAAGCAAGAGATATGTTTCTTATGAGATCTGATGCTGCTACATA 1980
QY 1981 GGGTCAAGTCAAGCTGACATTCAGAGACATTCACAGAAATGATCTAGACCTGACAG 2040
Db 1981 GGGTCAAGTCAAGCTGACATTCAGAGACATTCACAGAAATGATCTAGACCTGACAG 2040
QY 2041 AATATATATCCAAATGAGGAGAGATTTAAGTCTCTGCTTTTGGGCTGATTCACCGCAT 2100
Db 2041 AATATATATCCAAATGAGGAGAGATTTAAGTCTCTGCTTTTGGGCTGATTCACCGCAT 2100
QY 2101 GAAATGTTACAGAGGTTTGAATGAGGAGATCTGAAGATTTGCACTGACAGAGAGCT 2160
Db 2101 GAAATGTTACAGAGGTTTGAATGAGGAGATCTGAAGATTTGCACTGACAGAGAGCT 2160
QY 2161 GAAACTGCGAGCTTGGGTTGTTATGATTAACACTGTTGAACATCTGTTGACAGT 2220

Db 2161 GAAATCGCGACGCTGGGGTGGTTATGATATACAGTGTAACTATCTTACACAGT 2220
 Oy 2221 GTGACAGTCTGACGAGATGGAACCCGCTGCTGCTGGAAGCAACTCTGAGAGGCTGCT 2280
 Db 2221 GTGACAGTCTGACGAGATGGAACCCGCTGCTGCTGGAAGCAACTCTGAGAGGCTGCT 2280
 Oy 2281 TGTCTATCTGATTTGGTTCATCCAGAAAACAATGCTACTGATGTCAGAACTACACACA 2340
 Db 2281 TGTCTATCTGATTTGGTTCATCCAGAAAACAATGCTACTGATGTCAGAACTACACACA 2340
 Oy 2341 AGATACGAAGTTTCTGCTCCAGTCAGGGGTGGAATACTGGAAGGCTCTGTTCTTGA 2400
 Db 2341 AGATACGAAGTTTCTGCTCCAGTCAGGGGTGGAATACTGGAAGGCTCTGTTCTTGA 2400
 Oy 2401 TCATTA 2406
 Db 2401 TCATTA 2406

RESULT 3
 AY091075
 LOCUS
 DEFINITION Arabidopsis thaliana unknown protein (At5g42480) mRNA, complete cds.
 ACCESSION AY091075
 VERSION AY091075.1
 KEYWORDS GI:20259550
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE
 AUTHORS
 1 (bases 1 to 2637)
 Yamada, K., Banh, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M., Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamitani, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shim, P., Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.
 Arabidopsis Full Length cDNA Clones
 Unpublished
 2 (bases 1 to 2637)
 Yamada, K., Banh, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M., Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamitani, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shim, P., Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.
 Direct Submission
 Submitted (21-MAR-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA
 RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAPL cDNAs (RAPL cDNA: 'RIKEN Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamitani, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

TITLE
 JOURNAL
 COMMENT
 The Salk, Stanford, PGSC (SSP) Consortium members carried out the sequencing and annotation of the RAPL cDNAs: Yamada, K., Banh, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M., Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamitani, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Shim, P., Southwick, A., Davis, R.W., Ecker, J.R. and Theologis, A.
 Yamada, K. (SSP/PGSC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGSC)

contributed equally to this work as Pis.
 Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.
 FEATURES
 Source
 1. 2637
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /chromosome="5"
 /clone="RAPL09-76-G11 (R19395)"
 /ecotype="Columbia"
 /note="This clone is in a modified pBluescript vector (pLC-1) as a BamHI/XhoI insert."
 1..2637
 /gene="At5g42480"
 1..113
 /gene="At5g42480"
 114..2519
 /gene="At5g42480"
 /codon_start=1
 /evidence=experimental
 /product="unknown protein"
 /protein_id="AA013895.1"
 /db_xref="GI:20259551"
 /translation="MEALSHVIGIGLSPPOLCRLPATTKLRSHNTSTTICSAKMDRLSDPNTSDSSSSPATATTATVAVSPSIRPBRHVIPIDFVQVLAQTHFLTDGTRAPBARSPKPOGFSDDALISRQIIQAACETLSNRSRREYEGLLDDEATVITVPNDKVPALCVLQEGETIIVKVGSLAKERLPKFKDQVIVMLAPLQVSDVADAAKRLNGISGRNLIWSVGGSALVGLTRKFMNEAFIRMVAEQVDLFPATPSINPASFVEYVALVAQAFIGKKFLLQDSDQFOOLQAKNAMEIPAMLYDRNMWEIDFGIERGICALLIGKVDCEKRWLGLDSDQYRNPATVEFVLENSRDNDDRLCLLETWTLAGVYPPRPDKKFKGDYDDPMVLSYLERVVGSSPLAALAMARIGAEHYKASMDLQVFPSPRITDRNSAEKPDQETVFSVDPGNNVGRDGEPPVFLAEARPSNEFTNDYALRAGVSESSVDETVEMSAVMLKEASVXIIAAGVATGISISFQRYFPLKSSSFQKDMVSMESDVATIGSRADDSALPRMDARTLENI VSKKQIKSLAFGPDHRIEMPEVLVDGMLKITMDRAETQGLVYDYTLTKLSVDSVTVAADQTRALVEATLEBSACLSDLVHPENNATVIRYTTTVEYFWSGKRTIGSVLAS"
 2520..2637
 /gene="At5g42480"

gene
 5'UTR
 CDS

ORIGIN
 3'UTR
 Query Match 99.7%; Score 2399.6; DB 15; Length 2637;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 2402; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy 1 ATGGAAGCTCGAAGTCAAGTGGCAATGGTCTCCCAATTCATTAATGCGAATTACCA 60
 Db 114 ATGGAAGCTCGAAGTCAAGTGGCAATGGTCTCCCAATTCATTAATGCGAATTACCA 173
 Oy 61 CCGGCGACGACAAAGCTCCGACGTAGCCACAACACTTACAACTATCTGCTCGGCACG 120
 Db 174 CCGGCGACGACAAAGCTCCGACGTAGCCACAACACTTACAACTATCTGCTCGGCACG 233
 Oy 121 AAATGGGCGGACGCTTCTCTCCGACTTCAATTCACCTCGAATTCCTCTCCCTCCG 180
 Db 234 AAATGGGCGGACGCTTCTCTCCGACTTCAATTCACCTCGAATTCCTCTCCCTCCG 293
 Oy 181 TTCGCAACCGGCAACACACCGCACTCTGCTCTCTGACCAATCTATTGATGATGCC 240
 Db 294 TTGCGCAACCGGCAACACACCGCACTCTGCTCTCTGACCAATCTATTGATGATGCC 353
 Oy 241 GAACGCCAGTCCCAATCCCAATGATTTCTTACCAAGATTAAGAGCTCAAAACATTTG 300
 Db 354 GAACGCCAGTCCCAATCCCAATGATTTCTTACCAAGATTAAGAGCTCAAAACATTTG 413
 Oy 301 TTAACCGATGGAATCAGAGAGCAATGAGCTAGGGTTTGAACCGCGCAATTTGGGT 360
 Db 414 TTAACCGATGGAATCAGAGAGCAATGAGCTAGGGTTTGAACCGCGCAATTTGGGT 473
 Oy 361 TTCAGCAGCAGCGCTTATCAGCGGAGACAGATTTCTTCAAGCTGTGTCGAAACTCTG 420

D	474	TTGACGACGACGCTTTAAATCAGCCGAGACAGATTCTTCAAGCTGCTTGGAACCTG	533
Q	421	TCATACTCGGCTTAGAAGAGATACAAATGAAGTCTCTTGATGATGAAGAACTACA	480
D	534	TCCTATCTCGGCTTAGAAGAGATACAAATGAAGTCTCTTGATGATGAAGAACTACA	593
Q	481	GTATCATGTATGCTTGGGATGAAGTCTCTGAGGCTCTCTGTATATGCAAGAAAGT	540
D	594	GTATCATGTATGCTTGGGATGAAGTCTCTGAGGCTCTCTGTATATGCAAGAAAGT	653
Q	541	GGTGAAGCTGAGATAGTCTTCCGCTTGATAGGCTCTGCTTAAGAGAGGTTCCCTAAG	600
D	654	GGTGAAGCTGAGATAGTCTTCCGCTTGATAGGCTCTGCTTAAGAGAGGTTCCCTAAG	713
Q	601	TGCTTTAAGCAAGATGATGTTTATGATTAAGCGCTGCTTCTGATGCTGAGGAGAT	660
D	714	TGCTTTAAGCAAGATGATGTTTATGATTAAGCGCTGCTTCTGATGCTGAGGAGAT	773
Q	661	GCTATGACATTTGATCCACTGATTTTATATCTGTTATGATGATTTGTTGAAGAACTTG	720
D	774	GCTATGACATTTGATCCACTGATTTTATATCTGTTATGATGATTTGTTGAAGAACTTG	833
Q	721	AAAGCTTTTACAGAGAGAGAGCAAGTACCTTGACCCGATTTTACGTGCAAAATTTGAT	780
D	834	AAAGCTTTTACAGAGAGAGAGCAAGTACCTTGACCCGATTTTACGTGCAAAATTTGAT	893
Q	781	GAGACTTTGGAAGAGATCACTCCGCTTATGCTTTGAGGACTTGGCTTACCGCTTGGT	840
D	894	GAGACTTTGGAAGAGATCACTCCGCTTATGCTTTGAGGACTTGGCTTACCGCTTGGT	953
Q	841	GATGATTAACGCTGCGAAGAACTAAATGTTTAAAGCGGTGTCGGAATATTTTGCTCT	900
D	954	GATGATTAACGCTGCGAAGAACTAAATGTTTAAAGCGGTGTCGGAATATTTTGCTCT	1013
Q	901	GTTGAGAGAGGTGAGCATCAGCTCTTGTGGGGTTTGAACCCGTGAAGATTTATGAT	960
D	1014	GTTGAGAGAGGTGAGCATCAGCTCTTGTGGGGTTTGAACCCGTGAAGATTTATGAT	1073
Q	961	GAGGCTTTTTCAGATGACAGCTGCTGAGCAGGTTGATCTTTTGTAGTACCCCAAGC	1020
D	1074	GAGGCTTTTTCAGATGACAGCTGCTGAGCAGGTTGATCTTTTGTAGTACCCCAAGC	1133
Q	1021	AAATATTCACAGAGATCATTTGAAGTTTACAGAGTTGCACTTGTCTTGTGGCTCAAGCT	1080
D	1134	AAATATTCACAGAGATCATTTGAAGTTTACAGAGTTGCACTTGTCTTGTGGCTCAAGCT	1193
Q	1081	TTTATTTGGTAAAGAACCACTTTTACAGATGCTGATTAAGCAATTCACAGCACTTCAAG	1140
D	1194	TTTATTTGGTAAAGAACCACTTTTACAGATGCTGATTAAGCAATTCACAGCACTTCAAG	1253
Q	1141	CAGGCTAAGGTAATGCTATGAGATTCCTGCAATGTTGATGATPACACGGAATATTTGG	1200
D	1254	CAGGCTAAGGTAATGCTATGAGATTCCTGCAATGTTGATGATPACACGGAATATTTGG	1313
Q	1201	GAGATAGACTTCGCTCTAGAAAGGGACTCTGTGCACTGCTTATAGGCAAAAGTTGATGA	1260
D	1314	GAGATAGACTTCGCTCTAGAAAGGGACTCTGTGCACTGCTTATAGGCAAAAGTTGATGA	1373
Q	1261	TGCGGTATGTGTTGGGCTTAGACAGTGAAGATTCACATTTAGAAATCCACGCTATTTGTG	1320
D	1374	TGCGGTATGTGTTGGGCTTAGACAGTGAAGATTCACATTTAGAAATCCACGCTATTTGTG	1433
Q	1321	GAGTTTGTGTTGAGAAATTCAAATCGTATGACAAATGATGATCCCTGGGACTATGCAAA	1380
D	1434	GAGTTTGTGTTGAGAAATTCAAATCGTATGACAAATGATGATCCCTGGGACTATGCAAA	1493
Q	1381	TTGTTGGAACCTGTGTGCGAGGGTGTCTTTCTTAGTTCAGAGACACCAAAAGATPAA	1440
D	1494	TTGTTGGAACCTGTGTGCGAGGGTGTCTTTCTTAGTTCAGAGACACCAAAAGATPAA	1553
Q	1441	AAATTTAACTCGGGGACTACTATGATGATCTTATGTTTGAATTCCTTGAAAAGATG	1500
D	1554	AAATTTAACTCGGGGACTACTATGATGATCTTATGTTTGAATTCCTTGAAAAGATG	1613
Q	1501	GAGTATGTTCAAGGTTCTCCCTTAGCTGCTGCAACTATGGAAGAGATTGAGCCGAG	1560
D	1614	GAGTATGTTCAAGGTTCTCCCTTAGCTGCTGCAACTATGGAAGAGATTGAGCCGAG	1673
Q	1561	CATGTGAAGCTATGTGCTATGACAGGCACTGACAGAAAGTTTTCCTCCGCTATACAGAT	1620
D	1674	CATGTGAAGCTATGTGCTATGACAGGCACTGACAGAAAGTTTTCCTCCGCTATACAGAT	1733
Q	1621	AGAACTCCGCTGAACCCAGAGATGTCAGAGACAGATGTTTATGATGATCTCTTTGGT	1680
D	1734	AGAACTCCGCTGAACCCAGAGATGTCAGAGACAGATGTTTATGATGATCTCTTTGGT	1793
Q	1681	AACAAATGAGCCGTATGATGATGAGCCCTGTGCTTTATTTACAGAAAGCTTGAAGCCCT	1740
D	1794	AACAAATGAGCCGTATGATGATGAGCCCTGTGCTTTATTTACAGAAAGCTTGAAGCCCT	1853
Q	1741	GAAAACCTTTGAAACTATATGATTAATGCAATTCAGACTGGGGTCTCAAGAGTACGTTGAT	1800
D	1854	GAAAACCTTTGAAACTATATGATTAATGCAATTCAGACTGGGGTCTCAAGAGTACGTTGAT	1913
Q	1801	GAAACTATCTGTTGAAATGTCCGTTGCTGATATGTTTAAAGAGGCAAGTGAAGATCTTA	1860
D	1914	GAAACTATCTGTTGAAATGTCCGTTGCTGATATGTTTAAAGAGGCAAGTGAAGATCTTA	1973
Q	1861	GCTGCTGTGTGGCAATTTGACTGATTTTCACTGCTTACGACAGAGATTTTCTTAAAGC	1920
D	1974	GCTGCTGTGTGGCAATTTGACTGATTTTCACTGCTTACGACAGAGATTTTCTTAAAGC	2033
Q	1921	AGCTCATCTTTTCAACGCAAGATATATGTTTCTTATATGGAATCTGATGTGCTACATA	1980
D	2034	AGCTCATCTTTTCAACGCAAGATATATGTTTCTTATATGGAATCTGATGTGCTACATA	2093
Q	1981	GGGTCACTGACAGCTAGCAAGTTGCAAGCACTTCCAGAAATGATGCTAGACCTGAGAG	2040
D	2094	GGGTCACTGACAGCTAGCAAGTTGCAAGCACTTCCAGAAATGATGCTAGACCTGAGAG	2153
Q	2041	AATATAGTATCCAAAGTGCAGAAAGATTAAGTCTTGAGCTTTTGGGCTGATCAACGATA	2100
D	2154	AATATAGTATCCAAAGTGCAGAAAGATTAAGTCTTGAGCTTTTGGGCTGATCAACGATA	2213
Q	2101	GAAATGTTTACCAAGGTTTGTGATGAGGCGAATGCTGAAGTTTGAAGCTGAACAGACGT	2160
D	2214	GAAATGTTTACCAAGGTTTGTGATGAGGCGAATGCTGAAGTTTGAAGCTGAACAGACGT	2273
Q	2161	GAAATGCGGAGCTGGGCTGTTTATGATTAATCACTGTTGAACATATCTGTTGACAGT	2220
D	2274	GAAATGCGGAGCTGGGCTGTTTATGATTAATCACTGTTGAACATATCTGTTGACAGT	2333
Q	2221	GTGACAGTCTCAGACAGATGGAACCCGCTCTGTGTGAACCACTGAGAGAGTCTGCT	2280
D	2334	GTGACAGTCTCAGACAGATGGAACCCGCTCTGTGTGAACCACTGAGAGAGTCTGCT	2393
Q	2281	TGTCTATCTGATTTGTTGTTTCAAGAAAACATGCTATGATGTCAGAACCTTACACACA	2340
D	2394	TGTCTATCTGATTTGTTGTTTCAAGAAAACATGCTATGATGTCAGAACCTTACACACA	2453
Q	2341	AGATACGAAGTTTTCGTGTCACAGTCAAGGTGGAAGAAATCACTGAAGCTCTGTTTGGCA	2400
D	2454	AGATACGAAGTTTTCGTGTCACAGTCAAGGTGGAAGAAATCACTGAAGCTCTGTTTGGCA	2513
Q	2401	TCATPAA 2406	
D	2514	TCATPAA 2519	
RESULT 4			
AY221468 3668 bp DNA linear PLN 05-AUG-2003			
LOCUS Arabidopsis thaliana division protein (At6g6) gene, complete cds;			
DEFINITION nuclear gene for chloroplast product.			
ACCESSION AY221468			


```

VERSION      AY221468.1  GI:33436338
KEYWORDS
SOURCE       Arabidopsis thaliana (thale cress)
ORGANISM     Arabidopsis thaliana
              Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Rosids; eudotsids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE    1 (bases 1 to 3668)
AUTHORS      Vitina,S., Froehlich,J.E., Koksharova,O., Pyke,K.A., Van Ert,H. and
              Osteryoung,K.W.
TITLE        ARC6 is a J-Domain Plastid Division Protein and an Evolutionary
              Descendant of the Cyanobacterial Cell Division Protein Ftn2
JOURNAL      Plant Cell 15 (8), 1918-1933 (2003)
PUBMED       12897262
REFERENCES   2 (bases 1 to 3668)
AUTHORS      Vitina,S., Koksharova,O., van Ert,H., Froehlich,J.E. and
              Osteryoung,K.W.
TITLE        Direct Submission
JOURNAL      Submitted (17-JUN-2003) Department of Plant Biology, Michigan State
              University, 166 Plant Biology Building, East Lansing, MI 48824, USA
              Location/Qualifiers
FEATURES
source       1..3668
              /organism="Arabidopsis thaliana"
              /mol_type="genomic DNA"
              /db_xref="taxon:3702"
              /ecotype="Wassilewskija"
              /date="1998-01-01"
              /note="ARC6"
              /product="division protein"
              /protein_id="AA018645.1"
              /codon_start=1
              /note="inner envelope membrane-localized; similar to
              cyanobacterial cell division protein sli10.69 of
              Synchocystis; contains J-domain"
              /protein_id="AA018645.1"
              /db_xref="GI:33436338"
              /translation="MEALSHVGIQLSPQLRLPPATYTKLRSHNTSTTICGSKMAD
              RLSDFNFTSDSSSSPATATTTATLVLPISIDPEHHVPIPIFVLAQTFLT
              DGIRRAFAARVSKPQFSDALISROILOAACEISNPSREVEGLDDEAT
              VITPMDKVPALCVLOEGGETEIVLVGALLKRLPKRKQVTVMLALPIDS
              RDMALDPPDITGYEYEBALKILOEGGASLAPDLAQIDTELETTPRYVLELG
              LPGLDIAAKRLNGISVGNILMSVGGGASALVGLTRKPMNEAFIRMTAAEQVDL
              FVATPSNIPASFEYVALALVAQAFIGKPHLQADADKQFQLOQAKVAMEIPAM
              LYDTRNMEIDPGLERGLCALLIGKVDKRMGLGSDSOSVNPALIVEFVLENSRD
              DNDLPLGKLEETLAVGVPPRPDTRDKKPKLGDYDDPMVLSYERVAVGOSPL
              AAATMARIAGEHVASAMOLQKVPKRYDRKASAPKDOETVPSVDPRGNVGRD
              GERVETAEAVRPSNERTNDYAIKAGVSSSVDETTVMVADMLKASVYKILAGV
              AIGLSLPSQKFLPKSSSSFFQKDMVSSMESDVATIGVRADDSALPRMARTANT
              VSKQKISLAFKPDHRIEMLPEDVLDGMLIMTDRAPETQGLGVYDTLLKISVDS
              VTVSADGTRALVEATLIESACISDLVHPENNATDVRYTTRVEVFWMSKGMKITEGVS
              LAS"
ORIGIN
Query Match 71.7%; Score 1724; DB 15; Length 3668;
Best Local Similarity 79.2%; Pred. No. 0;
Matches 2406; Conservative 0; Mismatches 0; Indels 632; Gaps 5;
QY 1 ATGGAAGCTGAGTCAGTCAGTGGGCAATTGCTCTCCCATTCATTTATGCCGATTACA 60
DB 478 ATGGAAGCTTGAAGTCAGTCAGTGGGCAATTGCTCTCCCATTCATTTATGCCGATTACA 537
QY 61 CCGGGAAGCAAAAGCTCGAGCTAGCAACAACCTCTACAACTATCTGCTCGCAGC 120
DB 538 CCGGGAAGCAAAAGCTCGAGCTAGCAACAACCTCTACAACTATCTGCTCGCAGC 597
QY 121 AAATGGCCGACCGTCTTCTCCGACTTCAATTTCACTCCGATTCCTCTCTCTCC 180

```

```

DB 598 AAATGGCCGACCGTCTTCTCTCCGACTTCAATTTCACTCCGATTCCTCTCTCTCC 657
QY 181 TTGGCAGCCGACCAACCAACGCGCACTCTGCTCTGCGCAACATCTATTGATGCTCC 240
DB 658 TTGGCAGCCGACCAACCAACGCGCACTCTGCTCTGCGCAACATCTATTGATGCTCC 717
QY 241 GAACGCAAGTCCCAATCCCATTTGATTTCTACAGAGTATTAGAGCTCAACACATTTTC 300
DB 748 GAACGCAAGTCCCAATCCCATTTGATTTCTACAGAGTATTAGAGCTCAACACATTTTC 777
QY 301 TTAAACGATGATCAAGAAAGACATTTGAGAGTATGAGGTTTGAACCGCGCAATTCGGT 360
DB 778 TTAAACGATGATCAAGAAAGACATTTGAGAGTATGAGGTTTGAACCGCGCAATTCGGT 837
QY 361 TTCAGCGACAGCGCTTTATATGACCGGAGACAGATTTCTTCAAGCTCTTGGCAACTCTG 420
DB 838 TTCAGCGACAGCGCTTTATATGACCGGAGACAGATTTCTTCAAGCTCTTGGCAACTCTG 897
QY 421 TCTAATCTCGCTCTAGAGAGATGACATGAAAGGCTCTTCTGATGATGAAAGACTACA 480
DB 898 TCTAATCTCGCTCTAGAGAGATGACATGAAAGGCTCTTCTGATGATGAAAGACTACA 957
QY 481 GTCATCACTGATGCTTCTGGGAT----- 504
DB 958 GTCATCACTGATGCTTCTGGGATTAAGGTAATTCGATTCGGAATTAAGATTTCTTC 1017
QY 505 -----AAGTTCCTGGGCG 518
DB 1018 GTTTAAATTTCAATGATGATTAAGAAAGAACTTTATCTAGTAAGGTTCTGGGGC 1077
QY 519 TCTCTGNTATTTCAAGAAAGGTGTGATGACTGATAGTCTTCCGGTGTGAGAGCTCT 578
DB 1078 TCTCTGNTATTTCAAGAAAGGTGTGATGACTGATAGTCTTCCGGTGTGAGAGCTCT 1137
QY 579 GCTTAAGAGAGGTTGCTTAAGTCTTTAAGCAAGATGCTTTTATGATTAGCGCTTGC 638
DB 1138 GCTTAAGAGAGGTTGCTTAAGTCTTTAAGCAAGATGCTTTTATGATTAGCGCTTGC 1197
QY 639 GTTTCGATGCTCGAAGGATGCTAATGCAATGCAATTCACCTGATTTTATTAAGCTTA 698
DB 1198 GTTTCGATGCTCGAAGGATGCTAATGCAATGCAATTCACCTGATTTTATTAAGCTTA 1257
QY 699 TGAATTTGTTGAGAGGCTTTGAGAGCTTTAC----- 730
DB 1258 TGAATTTGTTGAGAGGCTTTGAGAGCTTTACAGGTATTGACTTGTGTTGTAATTTG 1317
QY 731 -----A 731
DB 1318 ACGAGCGTTGCTTTATTAAGAACTTTCTGATTTGATCTTTGTTATGAGTCTTGCTGA 1377
QY 732 GGAGGAAGCAAGTACGCTTTCGACCGGATTTACGTCACAAATTTGATGAGACTTTGGA 791
DB 1378 GGAGGAAGCAAGTACGCTTTCGACCGGATTTACGTCACAAATTTGATGAGACTTTGGA 1437
QY 792 AGAGATCACTCGGCTTATGCTCTTGAGAGCTACTTGAGCTTACCGCTTGATGATTAAGC 851
DB 1438 AGAGATCACTCGGCTTATGCTCTTGAGAGCTACTTGAGCTTACCGCTTGATGATTAAGC 1497
QY 852 TCGAAGAAAGCTAAATGCTTTAAGCGGTGTCGGAATATTTTGTGCTGTGAGAGAG 911
DB 1498 TCGAAGAAAGCTAAATGCTTTAAGCGGTGTCGGAATATTTTGTGCTGTGAGAGAG 1557
QY 912 TGGAGCATCAGCTCTTGTGTGGGGGTTTGAACCGGAGAAAGTTTATGATGAGAGCGTTTT 971
DB 1558 TGGAGCATCAGCTCTTGTGTGGGGGTTTGAACCGGAGAAAGTTTATGATGAGAGCGTTTT 1617
QY 972 ACGATGACAGCTGCTGAGC----- 991
DB 1618 ACGATGACAGCTGCTGAGGAGGATTAACGTTTATGATTAATTTTATTTTACGA 1677
QY 992 -----AAGTTGATCTTTTGTGA 1008

```


[illegible]

QY	1979	-----	1978
Db	2818	AMGATTAAACAAGTTGCTGAGTAATAATTCACATAATTATGCTGTGAATTTTGTGATCAAA	28777
QY	1979	-----	1978
Db	2878	CTGTAGACAGAAATGTAAATTTCACTCTCAACATTTCTGTTTAGAATAACGTAGGATTAG	2937
QY	1979	-----	1979
Db	2938	AGATTGCGCTTAGTGATGGGCTTTGTCACACTTTCTTCTCCTTGATTTTCTTTTCATTT	29977
QY	1980	AGGCTCAGTCAAGAGCTGACGATTGAGAAAGCACTTCCAGAAATGGAATGCTAGACTGAGA	2039
Db	2998	AGGCTCAGTCAAGAGCTGACGATTGAGAAAGCACTTCCAGAAATGGAATGCTAGACTGAGA	3057
QY	2040	GAATATGTATCCAAAGTGGCGAAGATTAAGTCTGAGCTTTTGGGCTGATTCACCGCAT	2099
Db	3058	GAATATGTATCCAAAGTGGCGAAGATTAAGTCTGAGCTTTTGGGCTGATTCACCGCAT	3117
QY	2100	AGAAATGTTACCGA-----	2113
Db	3118	AGAAATGTTACCGAAGGTGAGGGAATAATCTACATTCAATCAATTAATGTGTGAAAACGT	3177
QY	2114	-----	2128
Db	3178	TGCAATGATTATAGTCTGGGCTCTGTTGATTCTGTATTTATAGTGTGGATGGGC	3237
QY	2129	GAATGCTGAAGATTGGAATGGAACGACAGCTGAACTCGGCGAGCTGGTGGTTATG	2188
Db	3238	GAATGCTGAAGATTGGAATGGAACGACAGCTGAACTCGGCGAGCTGGTGGTTATG	3297
QY	2189	ATTATACACTGTTGAAACTATCTGTTGACAGTGAAGTCTCAGACAGATGAAGACCCGTG	2248
Db	3298	ATTATACACTGTTGAAACTATCTGTTGACAGTGAAGTCTCAGACAGATGAAGACCCGTG	3357
QY	2249	CTCTGTGGGAAGCACTCTGGAGAGTCTGTTGTTATCTGATGATTTGTTTCATCCAGAAA	2308
Db	3358	CTCTGTGGGAAGCAACTCTGGAGAGTCTGTTGTTATCTGATGATTTGTTTCATCCAGAAA	3417
QY	2309	ACATATGCTACTGATGTGAGAACTTACACACACAAAGATTTCTGTGCTCAAGTCAG	2358
Db	3418	ACATATGCTACTGATGTGAGAACTTACACACACAAAGATTTCTGTGCTCAAGTCAG	3477
QY	2369	GGTGAATAATCACTGAAGGCTCTGTTCTTGATCATATA	2406
Db	3478	GGTGAATAATCACTGAAGGCTCTGTTCTTGATCATATA	3515
RESULT 5			
LOCUS	AY221467	3664 bp	DNA linear
DEFINITION	Arabidopsis thaliana truncated division protein (arce6) gene, arce6-1 allele, complete cds; nuclear gene for chloroplast product.		PLN 05-AUG-2003
ACCESSION	AY221467		
VERSION	AY221467.1	GI:33436274	
KEYWORDS			
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Arabidopsis thaliana		
REFERENCE	1. (bases 1 to 3664)		
AUTHORS	Vitsha,S., Froehlich,J.E., Kokscharova,O., Pyke,K.A., Van Ert,H. and Osteryoung,K.W.		
TITLE	ARCE6 is a J-Domain Plastid Division Protein and an Evolutionary		
JOURNAL	Descendant of the Cyanobacterial Cell Division Protein Ftn2		
REFERENCE	Plant Cell 15 (8), 1918-1933 (2003)		
PUBMED	12897262		
AUTHORS	2. (bases 1 to 3664)		
	Vitsha,S., Kokscharova,O., van Ert,H., Froehlich,J.E. and Osteryoung,K.W.		

TITLE Direct Submission
JOURNAL Submitted (17-JAN-2003) Department of Plant Biology, Michigan State University, 166 Plant Biology Building, East Lansing, MI 48824, USA
FEATURES
source

1.3664
/organism="Arabidopsis thaliana"

/mol_type="genomic DNA"

/db_xref="taxon:3702"

/ecotype="Ws-2"

/note="genotype: arc6 mutant"

1.3664

/gene="arc6"

/allele="arc6-1"

join(478..984,1067..1291,1379..>1621)

/gene="arc6"

/allele="arc6-1"

/product="truncated division protein"

join(478..984,1067..1291,1379..1621)

/gene="arc6"

/note="truncation caused by premature stop codon"

/codon_start=1

/product="truncated division protein"

/protein_id="AAQ18644.1"

/db_xref="GI:3436275"

/translation="MEALSHVIGLSPOLCRLPPATYKLRSHNTSTTIGSASKMD
RLSDPFTSDSSSSPATATTTATVLPSPIDRPHVPIPIFYOVLAGOTFLT
DGIRAFARVSKPQSPDALSIRQIIOACETLSRREVRVGLDDEAT
VITVPMKRVKALCVLOEGSTRIVLAVGELAKERIPKSKQDVIVMLALDVS
RDMAALDPDFITGIEVEALKLQEGSASLAPDLAQIDETLEITTPRVLELIG
LPLGDDYAKRLINGISVRNLIWSVGGSASALVGLTRKEMNEAFI"

ORIGIN

Query Match 71.5%; Score 1720.8; DB 15; Length 3664;
Best Local Similarity 79.1%; Pred. No. 0;
Matches 2404; Conservative 0; Mismatches 2; Indels 632; Gaps 5;

```

Oy 1 ATGAAAGCTGTGAGTCAAGTCGGCAATTGGTCTCTCCCAATTCAAATTATGCGATTACA 60
Db 478 ATGAAAGCTGTGAGTCAAGTCGGCAATTGGTCTCTCCCAATTCAAATTATGCGATTACA 537
Oy 61 CCGGGAAGCAAAAGCTCGAGCTGACCAACACCTTACAACTATCTGCGCCGAC 120
Db 538 CCGGGAAGCAAAAGCTCGAGCTGACCAACACCTTACAACTATCTGCGCCGAC 597
Oy 121 AAATGGCGCAACCGCTTCTCTCGAATTCACTTCGATTCCTCCCTCCCTCC 180
Db 598 AAATGGCGCAACCGCTTCTCTCGAATTCACTTCGATTCCTCCCTCCCTCC 657
Oy 181 TTGCGCAACCGCAACCGCACTCTGCTCTCTGCAACATCTATGATGATCC 240
Db 658 TTGCGCAACCGCAACCGCACTCTGCTCTCTGCAACATCTATGATGATCC 717
Oy 241 GAAGCGCAAGTCCCATCCCATTTATTTCTACAGGATTAAGAGCTCAACACATTT 300
Db 718 GAAGCGCAAGTCCCATCCCATTTATTTCTACAGGATTAAGAGCTCAACACATTT 777
Oy 301 TTAACCGATGATGAAGAAGCATTTCGAAGCTAGAGTTTGAACCGCGCAATTGGGT 360
Db 778 TTAACCGATGATGAAGAAGCATTTCGAAGCTAGAGTTTGAACCGCGCAATTGGGT 837
Oy 361 TTACGCGAAGCAGCTTTTATATCAGCGGAGAGAGATTTCTTCAAGCTGCTTGGAAA 420
Db 838 TTACGCGAAGCAGCTTTTATATCAGCGGAGAGAGATTTCTTCAAGCTGCTTGGAAA 480
Oy 421 TCTATCTCGGTCTAGAGAGAGATCAATGAAGTCTTCTTGAATGAAGAAGCTACA 480
Db 898 TCTATCTCGGTCTAGAGAGAGATCAATGAAGTCTTCTTGAATGAAGAAGCTACA 957
Oy 481 GTCATCACTGATGTTCTTGGAT----- 504
Db 958 GTCATCACTGATGTTCTTGGATTAAGGTAATTTGGAATTAATAAGTTTCTTC 1017

```

```

Oy 505 -----AAGTTCTCGGGC 518
Db 1018 GTTTAATTCATGAATGATAAAGAAAGAACTTTATCTAGAAAGTTCCCGGGC 1077
Oy 519 TCTCTGTATTTGCAAGAGGTGTGAGACTGAGATATGTTCTTGGGTTGGAGGCTCT 578
Db 1078 TCTCTGTATTTGCAAGAGGTGTGAGACTGAGATATGTTCTTGGGTTGGAGGCTCT 1137
Oy 579 GCTTAAGGAGGTTGCTTAAGTGTGTTAAAGAAATGTTGTTTATGATGCGCTTGC 638
Db 1138 GCTTAAGGAGGTTGCTTAAGTGTGTTAAAGAAATGTTGTTTATGATGCGCTTGC 1197
Oy 639 GTTCTCGATGCTCGAGGAGATGCTATGAGATTCACCTGATTTATTTACTGTTA 698
Db 1198 GTTCTCGATGCTCGAGGAGATGCTATGAGATTCACCTGATTTATTTACTGTTA 1257
Oy 699 TGAGTTTGTGAGAAAGTTTGAAGCTTTAC----- 730
Db 1258 TGAGTTTGTGAGAAAGTTTGAAGCTTTTACAGGTAGTTGACTGCTTGTGTAATTG 1317
Oy 731 -----A 731
Db 1318 ACGAGCGTGGCTTTATTAAGAACTTCTGATTTGATCTTGTATTTGAGTCTGTGTA 1377
Oy 732 GAGAGAAAGCAAGTACCTTGCACCGGATTTACGTGCACAAAATGATGAGACTTTGGA 791
Db 1378 GAGAGAAAGCAAGTACCTTGCACCGGATTTACGTGCACAAAATGATGAGACTTTGGA 1437
Oy 792 AGAGTCACTCCGGGTTATGCTTGAAGCTACTGGCTTACCCTTGGTATGATTAAGC 851
Db 1438 AGAGTCACTCCGGGTTATGCTTGAAGCTACTGGCTTACCCTTGGTATGATTAAGC 1497
Oy 852 TGCGAAAAGCTAATGTTTACGGGATGTCGGAATTTTGTGTCGTGAGAGAG 911
Db 1498 TGCGAAAAGCTAATGTTTACGGGATGTCGGAATTTTGTGTCGTGAGAGAG 1557
Oy 912 TGAGCATCAGCTTGTGAGGAGGTTTGAACCGGTGAGAGTTTATGATGAGGCGTTT 971
Db 1558 TGAGCATCAGCTTGTGAGGAGGTTTGAACCGGTGAGAGTTTATGATGAGGCGTTT 1617
Oy 972 ACGAATGACAGCTCTGAC----- 991
Db 1618 ACGAATGACAGCTCTGACAGGTATACGTTTATGATACCTTTTATTTCTTTAGCA 1677
Oy 992 -----AGTTGATCTTTTGA 1008
Db 1678 TGATATACCTTTAGGTTTCTTATTTATGATGTTGTGTTGTTGATCTTTTGA 1737
Oy 1009 GCTACCCCAAGCAATATTCAGCAGAGTCAATTTGAAGTTTACGAAGTTGACTTGCCT 1068
Db 1738 GCTACCCCAAGCAATATTCAGCAGAGTCAATTTGAAGTTTACGAAGTTGACTTGCCT 1797
Oy 1069 GTGGCTCAAGTTTATTTGTAAGAACCACTTTTACAGATGCTGATTAACCAATTC 1128
Db 1798 GTGGCTCAAGTTTATTTGTAAGAACCACTTTTACAGATGCTGATTAACCAATTC 1857
Oy 1129 CAGCAACTTCAGCAGGCTAAGTTATGCTATGAGATTCCTGCGAGTTTGTATGATACA 1188
Db 1858 CAGCAACTTCAGCAGGCTAAGTTATGCTATGAGATTCCTGCGAGTTTGTATGATACA 1917
Oy 1189 CGGAATATTTGAGAGATGACTTGGTCTAGAAAAGGGACTGTGACACTGCTTATAGGC 1248
Db 1918 CGGAATATTTGAGAGATGACTTGGTCTAGAAAAGGGACTGTGACACTGCTTATAGGC 1977
Oy 1249 AAAGTTGATGAATCCGTAATGTTGTTGGCTTGAACAGTGAAGATTCACATATAGGAAT 1308
Db 1978 AAAGTTGATGAATCCGTAATGTTGTTGGCTTGAACAGTGAAGATTCACATATAGGAAT 2037
Oy 1309 CCAGCTATTTGAGAGTTGTTTGAAGATTCAAATGATGATGATGATGATCTCCCT 1368
Db 2038 CCAGCTATTTGAGAGTTGTTTGAAGATTCAAATGATGATGATGATGATCTCCCT 2097
Oy 1369 GGAATATGCAAAATTTGTGAAAAGCTGTTGACAGGGGTTGCTTCTTCTAGGTTCAAGAC 1428

```

```

Db      2098  GGACATATGCAATTTGTTGAAACCTGTTGGCGAGGGGTTGCTTTCCTAGGTTGAGAGAC 2157
Qy      1429  ACCAAAGATTAATAATTTAACTCGGGGACTACTATGATGATCCATGTTTGGTTATG 1488
Db      2158  ACCAAAGATTAATAATTTAACTCGGGGACTACTATGATGATCCATGTTTGGTTATG 2217
Qy      1489  TTGGAAGAGTGAAGTGAAGTTCAGGGTTCCTTTAGCTGCTGCACTATGCGCAGG 1548
Db      2218  TTGGAAGAGTGAAGTGAAGTTCAGGGTTCCTTTAGCTGCTGCTGCACTATGCGCAGG 2277
Qy      1549  ATTGGAAGCCGAGCATGTTAAAGCTATGCTATGCGGCACTGCGAAGTTTCTTCC 1608
Db      2278  ATTGGAAGCCGAGCATGTTAAAGCTATGCTATGCGGCACTGCGAAGTTTCTTCC 2337
Qy      1609  CGCTATACAGATGAGAACTCGGCTGTAACCAAGATGTCAGAGAGACAGTGTATAGTGA 1668
Db      2338  CGCTATACAGATGAGAACTCGGCTGTAACCAAGATGTCAGAGAGACAGTGTATAGTGA 2397
Qy      1669  GATCCGTGTTGTAACAATGTAAGCCGTGATGCTGAGCCCTGCTTTATTTGCAAGAGCT 1728
Db      2398  GATCCGTGTTGTAACAATGTAAGCCGTGATGCTGAGCCCTGCTTTATTTGCAAGAGCT 2457
Qy      1729  GTAAGACCCCTCTGAAAACTTTGAAACTATATGCAATTCGAGCTGGGCTCTCAGAG 1788
Db      2458  GTAAGACCCCTCTGAAAACTTTGAAACTATATGCAATTCGAGCTGGGCTCTCAGAG 2517
Qy      1789  AGTAGCCGTTGATGAAACTACTGTTGAAATGTCGGTGTGATGATGTTAAAGAGCAGAT 1848
Db      2518  AGTAGCCGTTGATGAAACTACTGTTGAAATGTCGGTGTGATGATGTTAAAGAGCAGAT 2577
Qy      1849  GTGAAGATCTAGCTGCTGCTGCTGCGCAATGAGATGATTCGATTCAGCCAGAGATAT 1908
Db      2578  GTGAAGATCTAGCTGCTGCTGCTGCGCAATGAGATGATTCGATTCAGCCAGAGATAT 2637
Qy      1909  TTTCTTAAGACAGCTCATCTTTTCAAGCAGAGATATGTTTCTTCAATGAAATCTGAT 1968
Db      2638  TTTCTTAAGACAGCTCATCTTTTCAAGCAGAGATATGTTTCTTCAATGAAATCTGAT 2697
Qy      1969  GTCCCTACCA----- 1978
Db      2698  GTCCCTACCAAGATGATGATTAATGATGCAATTTTCATATATGCTGCTCAAAATA 2757
Qy      1979  ----- 1978
Db      2758  TGCTTGTGTTGAGCTAAGAACATAGTCCCACTTAATACATGTCGCAAAAGTTGATCC 2817
Qy      1979  ----- 1978
Db      2818  AAGATTAACAAGTTGCTGAGTAATTTTCACTAATTAATGCTGCTGAAATTTTGTATCAA 2877
Qy      1979  ----- 1978
Db      2878  CTGTAGACAGAAATGTAATTTCACTCTCAACATTTCTGTTAGAAATACAGATGATAG 2937
Qy      1979  ----- 1979
Db      2938  AGATTTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2997
Qy      1980  AGGCTGAGTCAAGCTGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 2039
Db      2998  AGGCTGAGTCAAGCTGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 3057
Qy      2040  GAATATAGTATCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 2099
Db      3058  GAATATAGTATCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 3117
Qy      2100  AGAAATGTTACAG----- 2113
Db      3118  AGAAATGTTACAGAGTGAGGGAATAATCTACAAATTCATCAATGTTGTGAAGACTGT 3177
Qy      2114  -----AGTTTGTGATGGGC 2128

```

```

Db      3178  TGACATGATTAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3237
Qy      2129  GAATGCTGAAGATTTGAGTCAAGACAGACAGTGAATCTGCGACAGCTTGGTGGTTATG 2188
Db      3238  GAATGCTGAAGATTTGAGTCAAGACAGACAGTGAATCTGCGACAGCTTGGTGGTTATG 3297
Qy      2189  ATTATACACTGTTGAAACATATCTGTTGACAGTGTGACAGTCTGACAGATGAGAACCCGTG 2248
Db      3298  ATTATACACTGTTGAAACATATCTGTTGACAGTGTGACAGTCTGACAGATGAGAACCCGTG 3357
Qy      2249  CTCGTGTGGAAGAACCTGCGAGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2308
Db      3358  CTCGTGTGGAAGAACCTGCGAGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3417
Qy      2309  ACAATGCTACTGATGTGACAACTCAACAACAAGATTCGAAGTTTCTGCTCCAGTCAAG 2368
Db      3418  ACAATGCTACTGATGTGACAACTCAACAACAAGATTCGAAGTTTCTGCTCCAGTCAAG 3477
Qy      2369  GGTGGAATACACTGAAGGCTCTGTTCTTGATCATATA 2406
Db      3478  GGTGGAATACACTGAAGGCTCTGTTCTTGATCATATA 3515

RESULT 6
AB016888
LOCUS
DEFINITION
Arabidopsis thaliana genomic DNA, chromosome 5, PL clone:MDH9.
ACCESSION
AB016888 BAO00015
VERSION
AB016888.1 GI:3449329
KEYWORDS
SOURCE
ORGANISM
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1
REFERENCE
AUTHORS
Aasamizu, E., Sato, S., Kaneko, T., Nakamura, Y., Kotani, H., Miyajima, N.
and Tabata, S.
TITLE
Structural analysis of Arabidopsis thaliana chromosome 5. VIII.
Sequence features of the regions of 1,081,958 bp covered by
seventeen physically assigned pl and TAC clones
JOURNAL
DNA Res. 5 (6), 379-391 (1998)
PUBMED
10048488
JOURNAL
2 (bases 1 to 85791)
REFERENCE
Nakamura, Y.
Direct Submission
Submitted (18-0UG-1998) Yasukazu Nakamura, Kazusa DNA Research
Institute, Department of Plant Gene Research, 1532-3, Yana,
Kisarazu, Chiba 292-0812, Japan (E-mail: ynakamu@kazusa.or.jp,
tel:81-438-52-3935, fax:81-438-52-3934)
ADDRESS
For the latest information on annotation of this clone, please see
http://www.kazusa.or.jp/kaos/cgi-bin/sgd.graph.cgi?c=MDH9
Genes with similarity to proteins in the databases are described in
'product' or 'note' qualifiers. Genes that have no significant
protein similarity are described as 'unknown protein'.
The software programs used to predict genes include: Gra11
(Informatics Group, Oak Ridge National Laboratory,
http://compbio.ornl.gov/Gra11-1.3//),
GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html),
NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of
Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and
SplicePredictor (Volker Brendel, Stanford University,
http://gremind.zool.jastate.edu/cgi-bin/sp.cgi).
Genes encoding tRNAs are predicted by tRNAscan-SE
(Sean Eddy, Washington University School of Medicine, St. Louis,
http://genome.wustl.edu/eddy/tRNAscan-SE/).
This sequence may not be the entire insert of this clone. It may be
shorter because we remove overlaps between neighboring submissions.
The 5' clone is K5014 and the 3' clone is K16E1.
location/Qualifiers
1..85791
/organism="Arabidopsis thaliana"
FEATURES
source

```

/mol_type="genomic DNA"
/db_xref="taxon:3702"
/chromosome="5"
/clone_1b="Witsei P1"
/ecotype="Columbia"
complement(59. .239)
/note="CDS is reported in Acc# AB023032
gene_id:K5J14.12
unknown protein"
/number=1
/evidence=not experimental
complement(join(2056. .2079,2278. .2393,2991. .3633))
/note="contains similarity to unknown protein
gene_id:MDH9.1
pir||T04864"
/pseudo
/codon_start=1
/evidence=not experimental
join(5739. .6138,6462. .6595)
/note="unnamed protein product; gene_id:MDH9.2
unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAB10474.1"
/db_xref="GI:9759469"
/translation="VWDEFAEEDRKKSNLSIAKNSVPLFERGRFYEEYAKRNE
RLRRKGEAVGAVGKPYNLYGVNHEPMTRKGTAKKSIKTVVSMSSSIPR
YSLRKMAREKRPPLPLNVAVSAMKSVTRKEERSLILCTACNTSSGLAPFRLLTGS
EPGLVHKRTTLPFSALV"
complement(join(6868. .7911,7989. .8414,8555. .8671,
8754. .9137))
/note="gene_id:MDH9.3"
/codon_start=1
/evidence=not experimental
/product="arm repeat containing protein"
/protein_id="BAB10475.1"
/db_xref="GI:9759470"
/translation="MDTDEATGDANDEEVTAEPRIHDEVDAVEIHEEVEKEDD
DCEGADVIVSEFLDQINGRYRTOQKCEFLVRLKILPLDEIRVYKAPPEBIM
LNRLEKVLAKKLETCSSNGSIYMALEDGEMTRPFSYIEKLRVYVKAPEBIM
SGDADEIDSLCKOKAKRRITDIDELAVDMVVFSEKTPRNASATIRLAKGLE
LQTDILKTEITAIQSLIQDKGSLNIETKQHIIEILNFKKLQGLEAVDILYQPVINK
AITKSTSLIPHEFLCPITLIMLDPVIAIGQYEEKSIQKMPDAGHKTCPRQEL
DHLSTAPNALKNILMOWCEKNKMFKIPEKVSPOKXDEBSLIVELASSOLEBO
RSRVOMKRLARENPEMVLINAGATPLIVOLISYPSGIOENAVTLLNLSTDEVN
KLLISNEGAIPNITIEILNGBRARENSAALFSLSMDEKVTIGLSNGIPRYVDIL
OHTLRGKDALITALFNLISNSANKRAIDGIVQPLNLLKDKNLGMIDALSTILL
LASHPEGQALGQSLFIEITLVEFLRQGTPEKCATSVLLEIGSNSSFLIALLOFGV
YEYLVETITSGINRAQRKANALIQISKEQI"
10888. .12579
/note="unnamed protein product; gb|AA07802.1
gene_id:MDH9.4
similar to unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAB10476.1"
/db_xref="GI:9759471"
/translation="WMISEKPLGVBESIRQDELVTTSRRLYKVSQULKKKIKTEV
VEDEIARGAVNCLISVGCRAVDGDEDEEHLNKRMSASABDKGMTTGTETRL
DCFSYGRERFWKNNRKYIADSGODYRKHVYLPDILMCLMLPLTSLNAHLVCK
KQSMANTQRFLOMRREGSFQTPMLFLFALKDQSGSDIGYDVSODKMRITEDLL
KGRFMYSVTSLIHEEIIYVGRSMDRNSFKSRGILVSPSIIKAMRKIASMRHARSLPI
VGATEVSEFSTQTKONRODRPHLSRVGSESVYEDPHLSVARQNRNADQNGTK
SHRLIRODLRLNRSKRFLVLIAGGTGIDEPDLSGEIYDSATNTWSEMRDPMGF
GVSQCGIICNGIIFYAVSNDLSQVDIERGVTITQTSPIPRVHEFPKIVSCNHL
FMLSVMCDGEGQIGRRNKAVRKLMEIDLVTWTEYSVHPDAPMDNATVYSDQNI
LMGEMKFIHQVLSFFVPCDILTEASWRVSNRQNKNTLSCMKTALHLH"
complement(join(18111. .18193,18281. .18535,18773. .19001,
19088. .19195,19307. .19447,19542. .19952,20129. .20245))
/note="unnamed protein product; gene_id:MDH9.6
similar to unknown protein
sp|P42251"
/codon_start=1
/evidence=not experimental
/protein_id="BAB10478.1"
/db_xref="GI:9759473"
/translation="MNQISFWCSSLTLTVYIAAEPTITRAFGSCANQASAPQZIMDA
INKEDPOLFLMGDNIYGDIRRLPLVIGKRTFEPWNRSPRVSSEEMKRLAKAK
ANPGYSRIQRKAVYIGWDDHDYGLNDAGKPFKRVNOKMLDPLNEPLDSPRQA
GVYASYTYGPPNRKAVYIADTRHYRPLSDSGSLIDGTQMDMLENLSGPRSEITII
GSSQVLSNLSATGPLPFYMESWGRPFKRLFKLADTRGDVIFISGVHGEIT
RYDQSVGPIYDVTSGVLOSVEKVRPRLRLPLADTRGVIPSTWYINDCKRKECT
YGCQNRPAISIDMANPVTTIKLETRVDNGDTVLGTNYSLSLQGGNSLKDTTCKG
SQRYCTLEIDPGRYRLAVLIYFTAVLAMLGLIGAVLAITKCVKCKVD"
complement(21540. .22097)
/note="unnamed protein product; contains similarity to
calmodulin
gene_id:MDH9.7"
/codon_start=1
/evidence=not experimental
/protein_id="BAB10479.1"
/db_xref="GI:9759474"
/translation="MTLAKNOKSLRLYKVSSEKSSRNLEDESRTSNSGSSS
LNVELTVPYMANSDGKISGEILQSCVSLIGALISRVVEVSTVDYDGGFID
FEELTKMEBGDSDEKREKIKARPMYVMEGSEFTTAASRLTSLRSGSCYVDAC
KVMIRGFDDNDGDLSTDEFVLNMR"
join(24090. .24289,24410. .24614,24701. .24945,25115. .25313,
25632. .25679,25821. .25928,26039. .26227,26314. .26613,
26730. .26831,27169. .27303,27487. .27711,27919. .28122,
28213. .28380,28646. .28795,28889. .29026,29204. .29431,
29519. .29671,29784. .29910,30115. .30241,30558. .30674,
30763. .30843,30932. .31033,31112. .31262,31338. .31429)
/note="gene_id:MDH9.8"
/codon_start=1
/evidence=not experimental
/product="piti1ysin"
/protein_id="BAB10480.1"
/db_xref="GI:9759475"
/translation="MASSSSSIFGVSPFSLAPFNSGDSRRSHYLDKSNKYAFNPS
SPRLTPRVRVEAPSLIPYNGLMAOQNSHKLKNIYVSGKEGTSLSGCRFCLT
CKNOAGIRRALPSAFVDRTPAFLSRLSLSLRKSIOYNATIGPDEPHAPAFND
GIYAEKDDLDLPEIDSLAEFLGELTSHRKHGQAKNGRIYILPNKVPNPF
EAMNEVAGSITDEBEDQGIAMHTEHAFIASSKREKULGTGASNATVDPHNVFHI
HSPLHTYDSDDLPFVSLDALNELAEFPKFLSVEKERAAIISELOMWTIEYRVC
QLQHLQSENKDLGRRPPIGABEQIKKMDVRIKRFHRRVFPANATYIYGDIDNIR
IVNIEAVFKNQJDNSTSPSPKQAGMANFLVPLKLAGLGTSTSEKNTAODS
KMLKREHARIPVEHNWSLPGTSVLDKPOIPEHLLONPAIMPKIPIYSKVOTG
DIANVIMKRTFLSALHRIYTRKSSNPPTSVEILDSGREGCYTTITLTPSPON
KQNAVYKAVQVEKRLKRGVTRGELTRYMDALKDSHLLAMDNDVSVNDLTVRES
DALSHVMDTQGHETVAAGVTVLLEAVTGAQVLEFISDPFRPAAPALVAVV
PTKVHWDGVEESDFNISPDRIIESVKSGLAPLEABEDELVEPKELQSQULKEITLQ
NPCEVPIDSGGLTKLHDKETGITOLRSLNGIIVANKYSTTESAGVRLVGGGRALB
TSQSKGAVVGVRTLSGGGVDSRQVDFCVNHLINSLSSTSEETFIABEPFTLR
DNGMOAFOILLVHVLERSVWLEDAFDARQOLYISYPSIISLBRATRAHLLMTAMNG
DERFVEPTKSLQSLNLESYKDAVMSGFVDNNMVSIVGDPSEBEIRCLIDLYGTAK
ASHDSAPGSEBPLFRQPYAGIQOVFLKIDTRGCATYIAGAPRNKGFYVGGDL
FQSVSKLPLVAHDGLLSLEQDLBGGDEILQKLLRAHPLFGVYTMGLIAELIINSRLFTT

CDS

VRDLSGLTYDVSFEILNFDRLNLMWYIVSVTSIPKRYKAVADACKSVLRGLHSNOIAP
RELIDRAKRTLLMRHEALSKSNAYMLNLALQASVPEKXELKEILSTEYASIED
IYLAVNQLRVEDSDLSYSCIGLAGAOGBEITVLSEEBEPDVFSGVAVGKRSMTIR
PTT"
Complement {join(33251..33326,33512..33663,33744..33797,
33888..33968,34143..34265,34359..34768,34873..35052,
35052..35052)}

Query Match 71.4%; Score 1717.6; DB 15; Length 85791;

Best Local Similarity 79.1%; Pred. No. 0;

Matches 2402; Conservative 0; Mismatches 4; Indels 632; Gaps 5;

```
OY 1 ATGGAAGCTTGAAGTCAAGTGGGCAATGGTCTCTCCCATTCATTTATGCCGATTACA 60
DB 64077 ATGGAAGCTTGAAGTCAAGTGGGCAATGGTCTCTCCCATTCATTTATGCCGATTACA 64136
OY 61 CCGGCGAGCAAAAGCTCGAGAGTAGCCAAACACTCTACAACTATCTGGCTCCGACAGC 120
DB 64137 CCGGCGAGCAAAAGCTCGAGAGTAGCCAAACACTCTACAACTATCTGGCTCCGACAGC 64196
OY 121 AAATGGGCGAAGCTCTTCTCTCGACTTCAATTTCACTCCGATTCCTCTCTCTCC 180
DB 64197 AAATGGGCGAAGCTCTTCTCTCGACTTCAATTTCACTCCGATTCCTCTCTCTCC 64256
OY 181 TTGCGCACCGCCACCAACACCGCCACTCTGCTCTCTGCGACCATTTATGATGTCTCC 240
DB 64257 TTGCGCACCGCCACCAACACCGCCACTCTGCTCTCTGCGACCATTTATGATGTCTCC 64316
OY 241 GAAGCGCAGTCCCGCATCCCATTTGATTTCTACAGGATTAAGAGCTCAAAACATTTTC 300
DB 64317 GAAGCGCAGTCCCGCATCCCATTTGATTTCTACAGGATTAAGAGCTCAAAACATTTTC 64376
OY 301 TTAACCGATGAATAGAAAGAGCATTTGAGAGGATTTGAAACCGCGCAATTCGAT 360
DB 64377 TTAACCGATGAATAGAAAGAGCATTTGAGAGGATTTGAAACCGCGCAATTCGAT 64436
OY 361 TTACCGACGAGCTTTTATACGCCGAGACAGATTTCTTCAAGCTCTTGGAAACTGTG 420
DB 64437 TTACCGACGAGCTTTTATACGCCGAGACAGATTTCTTCAAGCTCTTGGAAACTGTG 64496
OY 421 TCTAATCTCGGCTTACAGAGAGATACATGAAAGCTCTTCTGATGATGAAGAGCTACA 480
DB 64497 TCTAATCTCGGCTTACAGAGAGATACATGAAAGCTCTTCTGATGATGAAGAGCTACA 64556
OY 481 GTCATCACTGATGTTCTTGGAT----- 504
DB 64557 GTCATCACTGATGTTCTTGGATTAAGGTAATTTGATTCGAAATAAAGTTCTTC 64616
OY 505 -----AAGTTCTTGGGCGC 518
DB 64617 GTTTTAATTTCAATGAATGGAATTAAGAAAGAACTTTATCTAAGTGAAGGTTCTGATGC 64676
OY 519 TCTCTGTGATTAAGAAAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 578
DB 64677 TCTCTGTGATTAAGAAAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 64736
OY 579 GCTTAAGAGAGTGTCTTAAGTCTTTAAGCAAGATGTGTTTATGATTAAGCGCTTGC 638
DB 64737 GCTTAAGAGAGTGTCTTAAGTCTTTAAGCAAGATGTGTTTATGATTAAGCGCTTGC 64796
OY 639 GTTTCTCGATGCTTGAAGAGATGCTATGAGATGATCACTGATTTTATTAATGCTGTTA 698
DB 64797 GTTTCTCGATGCTTGAAGAGATGCTATGAGATGATCACTGATTTTATTAATGCTGTTA 64856
OY 699 TGAAGTTTGTGAAGAGCTTTGAAGCTTTTAC----- 730
DB 64857 TGAAGTTTGTGAAGAGCTTTGAAGCTTTTACAGGATGATTTGACTTTGGTAATTTG 64916
OY 731 -----A 731
DB 64917 ACGAGCGTTGGCTTATAAGAACTTTCTGATTTGATACCTTTGTTATGAGCTTGTGTA 64976
OY 732 GGAGGAAGAGAGAGTGAAGCTTTGACCGGATTTAGTGAACAATGATGAGACTTTGGA 791
DB 732 GGAGGAAGAGAGAGTGAAGCTTTGACCGGATTTAGTGAACAATGATGAGACTTTGGA 791
```

```
DB 64977 GGAGGAAGAGAGAGTGAAGCTTTGACCGGATTTAGTGAACAATGATGAGACTTTGGA 65036
OY 792 AGAGATCACTCCGCGTATGCTTTGAGACTACTTGGCTTACCGCTTGGTATGATTAACGC 851
DB 65037 AGAGATCACTCCGCGTATGCTTTGAGACTACTTGGCTTACCGCTTGGTATGATTAACGC 65096
OY 852 TGGCAAAAGACTAATGTTTAAAGCGGTGCGGAATATTTTGTGCTGTGTGAGAGAGG 911
DB 65097 TGGCAAAAGACTAATGTTTAAAGCGGTGCGGAATATTTTGTGCTGTGTGAGAGAGG 65156
OY 912 TGGAGCATCAAGCTTTGTGGGGGTTTGAACCGTGAAGAACTTTATGAAAGGCGTTTT 971
DB 65157 TGGAGCATCAAGCTTTGTGGGGGTTTGAACCGTGAAGAACTTTATGAAAGGCGTTTT 65216
OY 972 ACGAATGACAGCTGCTGAGC----- 991
DB 65217 ACGAATGACAGCTGAGAGAGATACATTTAATCTTTTAAATTTCTTTAGCA 65276
OY 992 -----AGTTGATCTTTTGTGA 1008
DB 65277 TGATATTAATTTAGTCTCATTTTATGATGATGTTGTGATGATGATGATGATGATGAT 65336
OY 1009 GCTACCCCAAGCAATATTCAGAGAGAGTCAATTTGAAGTTTACGAAGTGAAGCTTGCCT 1068
DB 65337 GCTACCCCAAGCAATATTCAGAGAGAGTCAATTTGAAGTTTACGAAGTGAAGCTTGCCT 65396
OY 1069 GTGGCTCAAGCTTTATTTGTAAGAAAGCACACCTTTTACAGATGCTGATTAAGCAATTC 1128
DB 65397 GTGGCTCAAGCTTTATTTGTAAGAAAGCACACCTTTTACAGATGCTGATTAAGCAATTC 65456
OY 1129 CAGCACTTACAGAGGCTAAGGATATGCTATGAGATTTCTGCGATGTTGATGATACA 1188
DB 65457 CAGCACTTACAGAGGCTAAGGATATGCTATGAGATTTCTGCGATGTTGATGATACA 65516
OY 1189 CGGAATTAATTTGGAGATGAGCTTCCGCTGAAGAAAGGGAATCTGTGACGCTTATAGGC 1248
DB 65517 CGGAATTAATTTGGAGATGAGCTTCCGCTGAAGAAAGGGAATCTGTGACGCTTATAGGC 65576
OY 1249 AAGTTGATGATATGCCGATATGATGATGATGATGATGATGATGATGATGATGATGAT 1308
DB 65577 AAGTTGATGATATGCCGATATGATGATGATGATGATGATGATGATGATGATGATGAT 65636
OY 1309 CCAAGCTATTTGAGATTTGTTTGGAGATTTCAATGCTGATGACATGATGATGATGATGAT 1368
DB 65637 CCAAGCTATTTGAGATTTGTTTGGAGATTTCAATGCTGATGACATGATGATGATGATGAT 65696
OY 1369 GCACTATGCAAAATTTGTTGGAACCTGTTGGCAGGGTGTCTTCTCTAGCTCAGAGAC 1428
DB 65697 GCACTATGCAAAATTTGTTGGAACCTGTTGGCAGGGTGTCTTCTCTAGCTCAGAGAC 65756
OY 1429 ACCAAGATTAATAATTTAACTCGGGGACTAATGATGATGATGATGATGATGATGATGAT 1488
DB 65757 ACCAAGATTAATAATTTAACTCGGGGACTAATGATGATGATGATGATGATGATGATGAT 65816
OY 1489 TTGGAAGAGTGAAGTGAAGTTCAAGGTTCTCTTTAGCTCTGCTGCACTATGCAAGG 1548
DB 65817 TTGGAAGAGTGAAGTGAAGTTCAAGGTTCTCTTTAGCTCTGCTGCACTATGCAAGG 65876
OY 1549 ATTGAGACCGAGATGGAAGCTAGTGTATGAGGCACTGAGAAAGTTTCTCTCC 1608
DB 65877 ATTGAGACCGAGATGGAAGCTAGTGTATGAGGCACTGAGAAAGTTTCTCTCC 65936
OY 1609 CGCTATACAGATAGAAACTCGGCTGAACCCAGAGATGTCAGAGACAGTGTATGATGA 1668
DB 65937 CGCTATACAGATAGAAACTCGGCTGAACCCAGAGATGTCAGAGACAGTGTATGATGA 65996
OY 1669 GATCTGTTGTAACAAATGTAAGCGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1728
DB 65997 GATCTGTTGTAACAAATGTAAGCGCTGATGATGATGATGATGATGATGATGATGATGATGAT 66056
OY 1729 GTAAGACCTCTGAAACCTTGAACCTAATGATTAAGCAATTTGAGAGCTGGGGTCTCAGAG 1788
DB 66057 GTAAGACCTCTGAAACCTTGAACCTAATGATTAAGCAATTTGAGAGCTGGGGTCTCAGAG 66116
```

OY	1789	AGTAGCGTTGATGAAACACACGTTGAAATGCCGTTGGTGAATGTTAAAGGAGGCAAGT	1848
Db	66117	AGTAGCGTTGATGAAACACACGTTGAAATGCCGTTGGTGAATGTTAAAGGAGGCAAGT	66117
OY	1849	GTCGAAGATCCAGCTGCTGTGTGGCAATTGAGCTGATTTCACTGTTCAAGCAGAGATAT	1908
Db	66117	GTCGAAGATCCAGCTGCTGTGTGGCAATTGAGCTGATTTCACTGTTCAAGCAGAGATAT	66233
OY	1909	TTTCTTAAAGCAGCTCATCTTTTCCAAAGCAGAGATATGTTTCTTCTATGAAATCTGAT	1968
Db	66237	TTTCTTAAAGCAGCTCATCTTTTCCAAAGCAGAGATATGTTTCTTCTATGAAATCTGAT	66237
OY	1969	GTGCGTACCA-----	1978
Db	66297	GTGCGTACCAATAGATATGATTTAAATGATGCAATTTCTATATATCGCATTTGCTCAAAATA	66355
OY	1979	-----	1978
Db	66357	TGCTTTGTTTGTGAGCTPAAGAAACATAGTTCACCTATATACATGTGCCAAAGTTGAC	66411
OY	1979	-----	1978
Db	66417	AAGATTAAACAAGTTGCTGAGTAAATTTCACTATATATATGCTGTAATTTTTGATCAA	66477
OY	1979	-----	1978
Db	66477	CTGTAGACAGAAATGTAAATTTCACTCTCAACATTTCTGTTTGAATAACTGAGATTAAG	66533
OY	1979	-----	1979
Db	66537	AGATTGCTTAGTGTGAGCTTTGTGCCAATTTTCTTCTGATTTTCTTTTCATGATTT	66599
OY	1980	AGGCTCAGTCAAGCTGTGACATTTCAAGAGCACTTCCAGAAATGGAATGCTAGACTGAGA	2039
Db	66597	AGGCTCAGTCAAGCTGTGACATTTCAAGAGCACTTCCAGAAATGGAATGCTAGACTGAGA	66655
OY	2040	GAATATAGTATCCAAAGTGGCAGAAAGATTAGTCTCTGGCTTTTGGGCGTGAATCCGCAT	2099
Db	66657	GAATATAGTATCCAAAGTGGCAGAAAGATTAGTCTCTGGCTTTTGGGCGTGAATCCGCAT	66711
OY	2100	AGAAATGTTTACAG-----	2113
Db	66717	AGAAATGTTTACAGAGGTGAGAGGAATATCTACAAATTCATCAATTTGTGTGAAACTGT	66777
OY	2114	-----	2128
Db	66777	TGGAATGATTAATAGTCTGTGAGCTTGTGATTTCTGTTATTTAATAGTTTGGATGGGC	66833
OY	2129	GAATGCTGAAGATTGGACTGACAGAGAGCTGAACTGGGCAAGCTGGGTTGGTTATG	2188
Db	66837	GAATGCTGAAGATTGGACTGACAGAGAGCTGAACTGGGCAAGCTGGGTTGGTTATG	66899
OY	2189	ATTATACACTGTGTAAACTATCTGTGTGACAGGTGTGACAGTCTCAGCAGATGAACCCGTG	2248
Db	66897	ATTATACACTGTGTAAACTATCTGTGTGACAGGTGTGACAGTCTCAGCAGATGAACCCGTG	66955
OY	2249	CTCTGTGTGAAGCAACTGTGAAGAGCTGTGTTGTCTATCTGATTTGGTTCAATCCAGAA	2308
Db	66957	CTCTGTGTGAAGCAACTGTGAAGAGCTGTGTTGTCTATCTGATTTGGTTCAATCCAGAA	67016
OY	2309	ACAATGTCACTGATGTGAGAACTTACACAACAAGATACGAAGTTTCTGTGTGCAAGTCAAG	2368
Db	67017	ACAATGTCACTGATGTGAGAACTTACACAACAAGATACGAAGTTTCTGTGTGCAAGTCAAG	67077
OY	2369	GCTGGAATAATCACTGAAGGCTGTGTTCTTGATCATATA	2406
Db	67077	GCTGGAATAATCACTGAAGGCTGTGTTCTTGATCATATA	67114

DEFINITION	Medicago truncatula clone mth2-155019, WORKING DRAFT SEQUENCE, 4 ordered pieces.
ACCESSION	AC158210
VERSION	AC158210.11 GI:71061528
KEYWORDS	HTG; HTGS PHASE2; HTGS DRAFT.
SOURCE	Medicago truncatula (barrel medic)
ORGANISM	Medicago truncatula
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eustroids I; Fabales; Fabaceae; Papilionoideae; Trifoliales; Medicago.
AUTHORS	1 (bases 1 to 133779) Lin,S., Dixon,R., May,G., Summer,L., Gonzales,B., Cook,D., Kim,D. and Roe,B.A.
TITLE	Medicago truncatula BAC Clone mth2-155019
REFERENCE	2 (bases 1 to 133779) Lin,S., Dixon,R., May,G., Summer,L., Gonzales,B., Cook,D., Kim,D. and Roe,B.A.
TITLE	Direct Submission
JOURNAL	Submitted (09-MAR-2005) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA
REFERENCE	3 (bases 1 to 133779) Lin,S., Dixon,R., May,G., Summer,L., Gonzales,B., Cook,D., Kim,D. and Roe,B.A.
AUTHORS	Direct Submission
TITLE	Submitted (22-JUL-2005) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA
JOURNAL	On Jul 22, 2005 this sequence version replaced gi:68342147.
COMMENT	----- Genome Center Center: Department Of Chemistry And Biochemistry The University Of Oklahoma Center code:UOKNOR ----- * NOTE: This is a 'working draft' sequence. It currently * consists of 4 contigs. Gaps between the contigs * are represented as runs of N. The order of the pieces * is believed to be correct as given, however the sizes * of the gaps between them are based on estimates that have * provided by the submittor. * This sequence will be replaced * by the finished sequence as soon as it is available and * the accession number will be preserved. * 1 43719: contig of 43719 bp in length * 43720 43819: gap of unknown length * 43820 83929: contig of 40110 bp in length * 83930 84029: gap of unknown length * 84030 106272: contig of 22243 bp in length * 106273 133779: gap of unknown length * 106373 133779: contig of 27407 bp in length. Location/Qualifiers 1. .133779 /organism="Medicago truncatula" /mol_type="genomic DNA" /db_xref="taxon:3880" /clone_1fb="mth2-155019" /clone_1fb="Medicago truncatula BAC library H2" 43720. .43819 /estimated_length=unknown 83930. .84029 /estimated_length=unknown 106273. .106372 /estimated_length=unknown
FEATURES	
source	
ORIGIN	
Query Match	10.1%; Score 243.6; DB 14; Length 133779;
Best Local Similarity	59.6%; Pzed No.2,4e-54;
Matches 535; Conservative	0; Mismatches 229; Indels 134; Gaps 2
683	ATTATTATACGCTTATAGATTGTGGAGGAAGCTTTGAAGCTTTTACAGAGGAAGAG 7422


```
Db 132882 ATTGCAATTCCTGGTCAATATGCTCTTCTGATGACTCACTGTATCATGTAGAGAGAGGGG 132941
QY 743 CAAGTAGCCTTGACACCGGATTTAGCGTCAAAATTGATGAGACTTTGAGAGATCATCTC 802
Db 132942 CAAAGCAGCCTTACACCGGATTTACAAACACAAATGATGAGACTTTGAGAGATPACCC 133001
QY 803 CGCGTTATGCTCTGAGAGCTACTTGGCTTACCGCTGTGATGATTTACCGTGAAGAAC 862
Db 133002 CACGTTGATATTAACATTTTAGCTTGCCCTCTTGATGATGAAACATGAGCGGAGAG 133061
QY 863 TAAATGCTTTAAGCGGTGCGGAAATATTTTGTGCTGTGAGAGAGTGAAGCATGAG 922
Db 133062 AAGAAAGGCTCCCAAGGTCCGCAACATTCGTGGCAGTTGAGAGTGAAGAGCAGCAG 133121
QY 923 CTCTGTTGGGGGTTTGACCGGTGAGAAATTATGATGAGAGCGTTTATGCAATGACAG 982
Db 133122 CAATTGCTGGAGTTTCACAGCTGAGATTTCAATGAAATGAGGCAATTCCTGCATATGAAG 133181
QY 983 CTGCTGAGC----- 991
Db 133182 CTGCCGAAACAGGTTCATAGAAATATATGTCATGTTTATTTTATTTTCTTCGCG 133241
QY 992 ----- 991
Db 133242 CCCCCCTAGTTTGTGTTATGAGAGTTTACATTTATCTCTTCAATTAATCTTTTATC 133301
QY 992 -----AGTTGATCTTTTGTAGCTACCCCAAGCAATATTCAGCAGAG 1034
Db 133302 TCCAAATCTTTTTTTAAAGTGAACCTTTTGTAGCAACCTTACCAATATTCAGCTGA 133361
QY 1035 GTCAATTTGAAGTTTACGAAGTTGACCTTGCTGTGCTCAAGCTTTTATTTGTTAGAA 1094
Db 133362 AAGTTTGGGCTTATGGGGTTGCACTTGATGTTGTCAGACMAAGCTTTGTTAGTAAAA 133421
QY 1095 GCCACACCTTTTACAGAGATGCTGATAGAACATTCAGCAACTTGCAGCGCTAAGTAT 1154
Db 133422 GCCACATCTTATCAAGATGCTGATTAATTTATTCATCAACTTCAACAACTTAAGTATAC 133481
QY 1155 GGCTATGAGATTCCTGCGATGTGTATGATAC-----ACGGAATATTTGGAGATAGA 1208
Db 133482 AAATATGAGAAATGCTCCTCTCTGTTTATTACTCCCATGAGATGAGAAAGAGAAAGTTGA 133541
QY 1209 CTTGCGTCTAGAAAGGGGACTCTGTGACCTGCTTATAGCAAAAGTATGATGATGCCGTAT 1268
Db 133542 TTTTCATTTAGAAAGGGGCTGTGTGCACTGCTGTGGGGAGCTTGATCAATGTCATC 133601
QY 1269 GTGCTTGGGCTTACAGATGAGAGATTCACAAATATAGAAATCCAGCTATTTGGAGTTGT 1328
Db 133602 ATGTTGGGTCTCGATAGTGAAGCTCACCTTATCGAAACCATCTATTTATAGACTTAT 133661
QY 1329 TTTGAGAAATTCAAATCGATGATGACAAATGATATCTCCCTGAGTATGCAAAATTTGGA 1388
Db 133662 TATGAAAACGAAAAGGTGATGAGAGACATGATCTTCTCTGAGCTCTGTAATTTGTTGA 133721
QY 1389 AACCTGTTGGCAGGGGTTGCTTTCTTAGTTCAGAGACACCAAGATTAATAAATTT 1446
Db 133722 GACATGTTGATGAGAGAGTGTTCCTCTAGGTTTAGAATACTTAAGAGACAACTTT 133779
```

RESULT 8
AP008208_011/c
WPCOMMENT
Sequence split into 360 fragments LOCUS AP008208 Accession AP008208

```
Fragment Name Begin End
AP008208_000 1 110000
AP008208_001 100001 210000
AP008208_002 200001 310000
AP008208_003 300001 410000
AP008208_004 400001 510000
AP008208_005 500001 610000
AP008208_006 600001 710000
AP008208_007 700001 810000
AP008208_008 800001 910000
```

```
AP008208_009 900001 1010000
AP008208_010 1000001 1110000
AP008208_011 1100001 1210000
AP008208_012 1200001 1310000
AP008208_013 1300001 1410000
AP008208_014 1400001 1510000
AP008208_015 1500001 1610000
AP008208_016 1600001 1710000
AP008208_017 1700001 1810000
AP008208_018 1800001 1910000
AP008208_019 1900001 2010000
AP008208_020 2000001 2110000
AP008208_021 2100001 2210000
AP008208_022 2200001 2310000
AP008208_023 2300001 2410000
AP008208_024 2400001 2510000
AP008208_025 2500001 2610000
AP008208_026 2600001 2710000
AP008208_027 2700001 2810000
AP008208_028 2800001 2910000
AP008208_029 2900001 3010000
AP008208_030 3000001 3110000
AP008208_031 3100001 3210000
AP008208_032 3200001 3310000
AP008208_033 3300001 3410000
AP008208_034 3400001 3510000
AP008208_035 3500001 3610000
AP008208_036 3600001 3710000
AP008208_037 3700001 3810000
AP008208_038 3800001 3910000
AP008208_039 3900001 4010000
AP008208_040 4000001 4110000
AP008208_041 4100001 4210000
AP008208_042 4200001 4310000
AP008208_043 4300001 4410000
AP008208_044 4400001 4510000
AP008208_045 4500001 4610000
AP008208_046 4600001 4710000
AP008208_047 4700001 4810000
AP008208_048 4800001 4910000
AP008208_049 4900001 5010000
AP008208_050 5000001 5110000
AP008208_051 5100001 5210000
AP008208_052 5200001 5310000
AP008208_053 5300001 5410000
AP008208_054 5400001 5510000
AP008208_055 5500001 5610000
AP008208_056 5600001 5710000
AP008208_057 5700001 5810000
AP008208_058 5800001 5910000
AP008208_059 5900001 6010000
AP008208_060 6000001 6110000
AP008208_061 6100001 6210000
AP008208_062 6200001 6310000
AP008208_063 6300001 6410000
AP008208_064 6400001 6510000
AP008208_065 6500001 6610000
AP008208_066 6600001 6710000
AP008208_067 6700001 6810000
AP008208_068 6800001 6910000
AP008208_069 6900001 7010000
AP008208_070 7000001 7110000
AP008208_071 7100001 7210000
AP008208_072 7200001 7310000
AP008208_073 7300001 7410000
AP008208_074 7400001 7510000
AP008208_075 7500001 7610000
AP008208_076 7600001 7710000
AP008208_077 7700001 7810000
AP008208_078 7800001 7910000
AP008208_079 7900001 8010000
AP008208_080 8000001 8110000
AP008208_081 8100001 8210000
```



```

Db      80900 ATGATATCTTTTCACAAAAACCGAATAGCATTCCTCCTGATGGTTTGAATTTACAT 80841
Oy      1054  GTTGACTGCTCTTGCTGCTCAAGCTTTTATGTGTAGAAAGCCACACTTTTACAGAT 1113
Db      80840 GTAGACTTGCATGATCGCTCAGCAATTTATAGTAAAGGCCAATTCTATCATGATG 80781
Oy      1114  GCTGTAAAGCAATTCACCAATTCAGAGGCTTAAAGTAAAGGATTCCTCGC 1173
Db      80780 GCGATGATCTTTTGAACAATCCAGAAAGTTCA-----C 80745
Oy      1174  ATGTTGTATGATACACGGAATTAATGGAGATAGACTTCGCTTAGAAGGGGACTCTGT 1233
Db      80744 ATAGGTTCTCATTAATGCTTATGATATATAGATGACCTTGAATGGAAGGGCACTTCGC 80685
Oy      1234  GCACGCTTATATAGCAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1293
Db      80684 TCATTGCTAGTCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 80625
Oy      1294  TCACAAATATAGCAATTCAGCTATTTGTGAGTTTGTGAGAAATTCAAATC--GTGAT 1350
Db      80624 TCACCAATACAGAACCCCAAAATTCAGAGTTTATGTAGCAACTTAGACATCAGTGAA 80565
Oy      1351  GACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1410
Db      80564 GAGAAATGATCTTCTTCAGGGCTGTGCAAGCTTTTGTGAGACTTGTCTTATCTTGAAGTT 80505
Oy      1411  TTTCTTGAAGTTCAGAGACCCAAAGATTAATACTCGGGGACTACTATGATGAT 1470
Db      80504 TTTCTTGAAGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 80445
Oy      1471  CCTATGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1530
Db      80444 CCAGAAATTTTAAAGTACTTAAAGATGAGAGGCTGTGCTTCTCATTTGCTGCT 80385
Oy      1531  GCTGCACTATGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1556
Db      80384 GCTGCTGCTATTTGCAAAACTTGCTGCT 80359

```

RESULT 9

```

AP004885/c 150462 bp DNA linear PLN 15-SEP-2004
LOCUS      Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 2,
DEFINITION PAC clone:p0575f10.
ACCESSION  AP004885
VERSION     AP004885.3 GI:41053009
KEYWORDS
SOURCE      Oryza sativa (japonica cultivar-group)
ORGANISM    Oryza sativa (japonica cultivar-group)
            Burkarya; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.

```

```

REFERENCE   1
AUTHORS    Sasaki, T., Matsumoto, T. and Yamamoto, K.
TITLE      Oryza sativa nipponbare (GAS) genomic DNA, chromosome 2, PAC
           clone:p0575f10
JOURNAL    Published Only in Database (2002)
AUTHORS    2 (bases 1 to 150462)
           Sasaki, T., Matsumoto, T. and Yamamoto, K.
REFERENCE   Direct Submission
           Submitted (20-MAR-2002) Takuji Sasaki, National Institute of
           Agricultural Sciences, Rice Genome Research Program, Kannondai
           2-1-2, Tsukuba, Ibaraki 305-8602, Japan
           (E-mail:tsasaki@ias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
           Tel:81-298-38-7441, Fax:81-298-38-7468)
           On Jan 21, 2004 this sequence version replaced gi:38142429.
           Genes were predicted from the integrated results of the following:
           GENSCAN (http://CCR-081.mit.edu/GENSCAN.html), FGENESH
           (http://opal.biology.gatech.edu/Genemark/), Glimmer
           (http://www.softberry.com/), GeneMark.hmm
           (http://opal.biology.gatech.edu/Glimmer/glmr_form.html), RiceHMM
           (http://rgp.dna.affrc.go.jp/RiceHMM/), SplicePredictor
           (http://bioinformatics.iastate.edu/cgi-bin/sp.cgi), sim4

```

FEATURES

SOURCE

(http://glabin.cse.psu.edu/html/docs/sim4.html), gap2 (http://www.tigr.org/software/glimmer/), BLASTN and BLASTX. The genomic sequence was searched against NCBI Nonredundant Protein database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at RGP or DDBJ. Protein homologies of the coding regions were searched against NCBI Nonredundant Protein database with BLASTP. ESTs represent the identified cDNA sequences using BLASTN with the corresponding DDBJ accession no. and RGP clone ID. Full-length cDNAs represent the identified cDNA sequences using BLASTN with the corresponding DDBJ accession no.

A gene with identity or significant homology to a protein is classified based on the protein name to indicate the homology level such as same name, 'putative-' and '-like protein'. A gene without significant homology to any protein but with full-length cDNA or EST homology (covering almost the entire length of partial sequence) is classified as an 'unknown' protein. A gene predicted by two or more gene prediction programs is classified as a 'hypothetical' protein according to IRGSP standard. A gene predicted by a single gene prediction program is also classified as a probable 'hypothetical' protein and is included as a miscellaneous feature of the sequence.

The orientation of the sequence is from 5' to 3' of the PAC clone. This sequence of p0575f10 clone has an overlap with Orl020.C02 (DDBJ: AP004078) clone at 5' end and with p0482f12 (DDBJ: AP005311) clone at 3' end. Detailed information on overlap and assembly quality together with annotation of this entry is available at <http://rgp.dna.affrc.go.jp/GenomeSeq.html>.

Location/Qualifiers

```

1..150462
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="2"
/clone="p0575f10"
/complement(4887..8112)
/gene="p0575f10.1"
/complement(join(4987..5412,5520..5679,5764..5914,6096..6247,6738..6859,7864..8112))
/gene="p0575f10.1"
/note="supported by full-length cDNA(s): AK106553"
/complement(join(5003..5412,5520..5679,5764..5914,6096..6247,6738..6859,7864..8084))
/gene="p0575f10.1"
/note="supported by full-length cDNA(s): AK073059"
/complement(join(5315..5412,5520..5679,5764..5914,6096..6247,6738..6859,7864..7876))
/gene="p0575f10.1"
/note="contains EST(s): AU070160(E61534)
/contains full-length cDNA(s): AK106553, AK073059"
/codon_start=1
/product="putative eukaryotic translation initiation factor 3, subunit 1 alpha, 35kDa"
/protein_id="BAD07941.1"
/db_xref="GI:41053010"
/translation="MEDWSEDPQAPVAPSAKAPLXKKADEVEDVDVKSWESEEEKPPEPVEKTAPKPSGKAAGKQOASSTSVVDEADLADPALKLRQRIIVEADFKSTTELFGKDGSEKSLDTFIPKSSSDPAVEKELANKLRPYEKSPFVWGLKNTVRKNSASLKADAKDISSSIAIANEKIKAKREAAAGKKQGAKKKHIEKKDDFLPGKNSFPDDPDERVDFW"
/complement(10186..13754)
/gene="p0575f10.2"
/complement(join(10186..10479,10759..10892,11127..11991,12089..12349,12515..12712,12843..13073,13455..13754))
/gene="p0575f10.2"
/note="start and end point are not identified"
/complement(join(10186..10479,10759..10892,11127..11991,12089..12349,12515..12712,12843..13073,13455..13754))
/gene="p0575f10.2"
/note="contains EST(s): AU058423(E51136), AU183658(E51136)
/codon_start=1
/product="plastid division protein precursor"
/protein_id="BAD07942.1"

```

CDS

gene

mRNA


```

Db      11715 TCACCATACAGAGACCCCAAAATCTAGATTATTGACCACTACATCAGTCA 11656
Qy      1351 GACATATGATCTCCCTGGAATATGCAAAATGTTGAAACCTGGTGAGGGCTGTC 1410
Db      11655 GAGATGATCTTCTTCGAGGGCTGCGAAGCTTTTGAGACTTGCTTATCTTTGAGTT 11596
Qy      1411 TTTCTAGGTTGAGACACCAAGATMAAAATTTAACTCGGAGACTATGATGAT 1470
Db      11595 TTTCTAGAGACGAGATCTCGGGCATGACATTGACATTGAGATTACATGATAT 11536
Qy      1471 CCTATGCTTTAGTACTTGAAGAGAGGAGTATGTTACAGGTTCTCTTACCTGCT 1530
Db      11535 CAGAAAGTTTAAGCTACTTAAGAAAGATGAGGGTGCTGCTTCTCATTTGGCTGCT 11476
Qy      1531 GCTGCACTATGCAAGATGTTGAGC 1556
Db      11475 GCTGCTGCTATTGCAAAACTTGCTGC 11450

RESULT 10
LOCUS      DQ022571          550 bp      DNA      linear      STS 28-MAY-2005
DEFINITION Beta vulgaris chromosome 1, sequence tagged site.
ACCESSION  DQ022571
VERSION     DQ022571.1 GI:66394762
KEYWORDS   STS.
SOURCE      Beta vulgaris
            Beta vulgaris
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            Caryophyllales; Amaranthaceae; Beta.
REFERENCE   1 (bases 1 to 550)
            Priesen,T.L., Weiland,J.J., Aasheim,M.L., Hunger,S., Borchardt,D.C.
            and Lewellen,R.T.
            Identification of a SCAR marker associated with Bm the Beet mosaic
            virus resistance gene on chromosome 1
            Unpublished
            2 (bases 1 to 550)
            Weiland,J.J. and Priesen,T.L.
            Direct Submission
            Submitted (29-APR-2005) Sugarbeet and Potato Research
            USDA-Agricultural Research Service, 1307 18th St. N, Fargo, ND
            58105, USA
FEATURES
    source          Location/Qualifiers
                    1..550
                        /organism="Beta vulgaris"
                        /mol_type="genomic DNA"
                        /cultiivar="C719"
                        /db_xref="taxon:161934"
                        /chromosome="1"
                        /note="fwd primer seq: GACTGAGTCTTAAGAAGCACTGT;
                        rev primer seq: GGAAGCATTTCACTCTTTATGCT;
                        fwd primer_name: Rbm05fwd; rev primer_name: Rbm05rev"
                    1..550
                        /note="SCAR marker associated with resistance to Beet
                        mosaic virus"

ORIGIN
Query Match      6.9%; Score 165.4; DB 10; Length 550;
Beet Local Similarity 74.6%; Pred.No.2.5e-33;
Matches 208; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

Qy      718 TTGAAGCTTTTACAGAGAGAGCACTGACCTTGACCGGATTTAGTCACAAAT 777
Db      102 TTCACTATTTTTGAAGAGAGGCTTCACTCTTGCGCCGATTTACAGACAGATT 161
Qy      778 GATGAGACTTTGAGAGATCACTCGCGTTATGCTTTGAGACTTCTTGACCGCTT 837
Db      162 GAGGAGACATTTGAGAGATCAAGACCCGCTTATGCTTGAGACTTTTGCTGCTC 221
Qy      838 GGTGATGATTTAGGCTCCGAAAGACATTAATGTTTAAAGCGGTGCGGAATTTTTCG 897
Db      222 AGTGATGATTTTCGAAAGAAAGATGAAGGCTTCACGCTGTACGCAATATATTATG 281

```

```

Qy      898 TCTGTTGAGAGAGGTGAGCATCAGCTCTTGTGGGGTTTGACCCCTGAGAAATTATG 957
Db      282 TCTGTTGAGAGAGGTGAGCGCAGCATATAGCTGTGTTTACTGCTGAGAAATTATG 341
Qy      958 AATGAGCGCTTTTACGAATGACAGCTGCTGAGCAGTT 996
Db      342 AATGAGCGCTTTTTCATATGACAGCAGCTGAGCAGTT 380

RESULT 11
LOCUS      AC155597          184752 bp      DNA      linear      HTG 25-JAN-2005
DEFINITION Zea mays strain B73 clone ZMMB8c0216K08, *** SEQUENCING IN PROGRESS
            **, 26 unordered pieces.
ACCESSION  AC155597
VERSION     AC155597.2 GI:58082456
KEYWORDS   HTG; HTGS_PHASE1.
SOURCE      Zea mays
            Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoidae; Andropogoneae; Zea.
            1 (bases 1 to 184752)
            Chan,A.P., Pertea,G., Zheng,L., Cheung,F., Lee,D., Koo,H.,
            Uteerback,T.R., Feldblyum,T.V., Rabinowicz,P., Fraser,C.M.,
            Schubert,K., Sam Miguel,P., Ma,J., Pontaroli,A.C., Rohlfing,T.,
            Budiman,A., Bedell,J.A., Lakey,N., Barbazuk,B., Bennetzen,J. and
            Quackenbush,J.
            Consortium for Maize Genomics - BAC skm sequencing and assembly
            Unpublished
            2 (bases 1 to 184752)
            Chan,A.P., Pertea,G., Zheng,L., Cheung,F., Lee,D., Koo,H.,
            Uteerback,T.R., Feldblyum,T.V., Rabinowicz,P., Fraser,C.M.,
            Schubert,K., Sam Miguel,P., Ma,J., Pontaroli,A.C., Rohlfing,T.,
            Budiman,A., Bedell,J.A., Lakey,N., Barbazuk,B., Bennetzen,J. and
            Quackenbush,J.
            Direct Submission
            Submitted (15-JAN-2005) The Institute for Genomic Research (TIGR),
            9712 Medical Center Dr, Rockville, MD 20850
            3 (bases 1 to 184752)
            Chan,A.P., Pertea,G., Zheng,L., Cheung,F., Lee,D., Koo,H.,
            Uteerback,T.R., Feldblyum,T.V., Rabinowicz,P., Fraser,C.M.,
            Schubert,K., Sam Miguel,P., Ma,J., Pontaroli,A.C., Rohlfing,T.,
            Budiman,A., Bedell,J.A., Lakey,N., Barbazuk,B., Bennetzen,J. and
            Quackenbush,J.
            Direct Submission
            Submitted (25-JAN-2005) The Institute for Genomic Research (TIGR),
            9712 Medical Center Dr, Rockville, MD 20850
            On Jan 25, 2005 this sequence version replaced gi:57863118.
            ----- Trace submission
            Center name: TIGR
            Seq_id: ZGGO
            ----- Project information
            Web site: http://www.tigr.org/cdb/cgi/maize/
            Contact: maize@tigr.org

* NOTE: This is a 'working draft' sequence. It currently
* consists of 26 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1
* 2210: contig of 2210 bp in length
* 2211
* 2310: gap of unknown length
* 2311
* 9941: contig of 7631 bp in length
* 9942
* 10041: gap of unknown length
* 10042
* 16745: contig of 6704 bp in length
* 16746
* 16845: gap of unknown length
* 16846
* 32394: contig of 15549 bp in length
* 32395
* 32494: gap of unknown length

```

```
* 32495 35494: contig of 3000 bp in length
* 35495 35594: gap of unknown length
* 35595 40051: contig of 4457 bp in length
* 40052 40151: gap of unknown length
* 40152 41379: contig of 1228 bp in length
* 41380 41479: gap of unknown length
* 41480 44517: contig of 3038 bp in length
* 44518 44564: gap of unknown length
* 44565 45664: gap of unknown length
* 45665 49440: contig of 3776 bp in length
* 49441 49541: gap of unknown length
* 49541 52698: contig of 3158 bp in length
* 52699 52798: gap of unknown length
* 52799 55964: contig of 3166 bp in length
* 55965 56064: gap of unknown length
* 56065 59236: contig of 3172 bp in length
* 59237 59336: gap of unknown length
* 59337 70052: contig of 10716 bp in length
* 70053 70152: gap of unknown length
* 70153 84603: contig of 14451 bp in length
* 84604 84703: gap of unknown length
* 84704 86418: contig of 1715 bp in length
* 86419 86518: gap of unknown length
* 86519 88069: contig of 1551 bp in length
* 88070 88169: gap of unknown length
* 88170 89661: contig of 1492 bp in length
* 89662 91357: gap of unknown length
* 91358 91457: contig of 1596 bp in length
* 91458 132727: contig of 41270 bp in length
* 132728 132827: gap of unknown length
* 132828 137468: contig of 4641 bp in length
* 137469 137568: gap of unknown length
* 137569 163843: contig of 26275 bp in length
* 163844 163943: gap of unknown length
* 163944 169858: contig of 5915 bp in length
* 169859 169958: gap of unknown length
* 169959 176100: contig of 6142 bp in length
* 176101 176200: gap of unknown length
* 176201 183079: contig of 6879 bp in length
* 183080 183179: gap of unknown length
* 183180 184752: contig of 1573 bp in length.
* Location/Qualifiers
* 1. 184752
  /organism="Zea mays"
  /mol_type="genomic DNA"
  /strain="B73"
  /db_xref="taxon:4577"
  /clone="ZM58BC0216K08"
  2211..2310
  /estimated_length=unknown
  gap 9942..10041
  /estimated_length=unknown
  gap 16746..16845
  /estimated_length=unknown
  gap 32395..32494
  /estimated_length=unknown
  gap 35495..35594
  /estimated_length=unknown
  gap 40052..40151
  /estimated_length=unknown
  gap 41380..41479
  /estimated_length=unknown
  gap 44518..44617
  /estimated_length=unknown
  gap 45565..45664
  /estimated_length=unknown
  gap 49441..49540
  /estimated_length=unknown
  gap 52699..52798
  /estimated_length=unknown
  gap 55965..56064
```

```
Query Match 6.8%; Score 163.4; DB 14; Length 184752;
Best Local Similarity 54.0%; Pred. No. 1.8e-32;
Matches 506; Conservative 0; Mismatches 296; Indels 135; Gaps 3;

ORIGIN
* 731 AGAGGAGGAGGAGAGTACGCTGACCGGATTACGCGCAAAATGATGAGCTTGG 790
  |||||
  /estimated_length=unknown
  gap 59237..59336
  /estimated_length=unknown
  gap 70053..70152
  /estimated_length=unknown
  gap 84604..84703
  /estimated_length=unknown
  gap 86419..86518
  /estimated_length=unknown
  gap 88070..88169
  /estimated_length=unknown
  gap 89662..89761
  /estimated_length=unknown
  gap 91358..91457
  /estimated_length=unknown
  gap 132728..132827
  /estimated_length=unknown
  gap 137469..137568
  /estimated_length=unknown
  gap 163844..163943
  /estimated_length=unknown
  gap 169859..169958
  /estimated_length=unknown
  gap 176101..176200
  /estimated_length=unknown
  gap 183080..183179
  /estimated_length=unknown

* 791 AAGAGTACACCGCGTATGCTTGGAGCTACTTGGCTTACCGCTTGATGATTAACG 850
  |||||
  /estimated_length=unknown
  gap 108996
  /estimated_length=unknown
  gap 183080..183179
  /estimated_length=unknown

* 851 CTGGAAAAAGCTAATGCTTTAAGCGGTGCGGAATTTTGTGCTGTTGGAGAG 910
  |||||
  /estimated_length=unknown
  gap 109056
  /estimated_length=unknown
  gap 183080..183179
  /estimated_length=unknown

* 911 GTGAGCATCAGCTCTTGTGGGGTTGACCGGCGAAGTTATGATGAGCGTTT 970
  |||||
  /estimated_length=unknown
  gap 109116
  /estimated_length=unknown
  gap 183080..183179
  /estimated_length=unknown

* 971 TACGAATGACAGCTGTCGAGGAGGTTGATCTTTTGTAGT----- 1011
  |||||
  /estimated_length=unknown
  gap 109176
  /estimated_length=unknown
  gap 183080..183179
  /estimated_length=unknown

* 1012 ----- 1011
  /estimated_length=unknown
  gap 109236
  /estimated_length=unknown
  gap 183080..183179
  /estimated_length=unknown

* 1022 -----ACCGAAGAAATTCGACGAGCATTTG 1042
  |||||
  /estimated_length=unknown
  gap 109296
  /estimated_length=unknown
  gap 183080..183179
  /estimated_length=unknown

* 1043 AAGTTACGAGTTGACTGCTCTGCTGAGCTTTTATTTAGTAAAGCCACACC 1102
  |||||
  /estimated_length=unknown
  gap 109356
  /estimated_length=unknown
  gap 183080..183179
  /estimated_length=unknown

* 109356 AGATCTATAGTGTGCACTTCCACATGTGTCAAGCAATTTGCAAGTCAAT 109415
  |||||
  /estimated_length=unknown
  gap 1103
  /estimated_length=unknown
  gap 109416
  /estimated_length=unknown
  gap 183080..183179
  /estimated_length=unknown

* 1103 TTTTACAGGATGCTGATTAAGCAATTCAGCAACTTCAGAGGCTTAAGTAAATG 1162
  |||||
  /estimated_length=unknown
  gap 109461
  /estimated_length=unknown
  gap 183080..183179
  /estimated_length=unknown

* 1163 AGATTCTGCGATGTTATGATACAGCAATTTAGGAGATGACTTGCGTCAAGAA 1222
  |||||
  /estimated_length=unknown
  gap 109462
  /estimated_length=unknown
  gap 183080..183179
  /estimated_length=unknown

* -----TATAGTTCTCATATCTTATGAAAAAGATGAGCTTCGATTAGAAA 109511
  |||||
  /estimated_length=unknown
  gap 109511
  /estimated_length=unknown
```


[illegible]

```

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
SOURCE
FEATURES
SOURCE
/organism="Rattus norvegicus"
/mol_type="genomic DNA"

```

/db_xref="taxon:10116"
/clone="CH230-228B13"
misc_feature
7131..8180
/note="wgs contig"
misc_feature
84069..8523
/note="wgs contig"
misc_feature
99455..102219
/note="wgs contig"
misc_feature
161841..163460
/note="wgs contig"
misc_feature
178218..179904
/note="wgs contig"
misc_feature
223846..225111
/note="wgs contig"
gap
227402..227501
/estimated_length=unknown
gap
228774..228873
/estimated_length=unknown
gap
231054..231153
/estimated_length=unknown
ORIGIN
Query Match 2.6%; Score 62; DB 14; Length 232974;
Best Local Similarity 43.3%; Pred.No. 7.9e-05;
Matches 290; Conservative 0; Mismatches 380; Indels 0; Gaps 0;
450 TGAAGGCTCTTGTGATGATGAAGAAGTACATCATCTGATGTTCTTGGGATAGGT 509
20297 TGAATTTGCTGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 20238
510 TCTTGAGGCTCTCTGTGTATTCAGAAAGGTGAGAGCTGAGATAGTTCTTGGGTTGG 569
20237 TGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 20118
570 TGAAGGCTCTGCTAAGAGAGAGTTGCCATAGCTGTTAAGCAAGATGCTTTAGTTAT 629
20177 TGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 20118
630 GGGCGTTCGCTTCGATGATTCAGAGGATGCTATGAGCATTCAGTCACTGATTTAT 689
20117 TGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 20058
690 TACTGCTTATGATTTGTTGAGAAAGCTTTGAAGCTTTTACAGAGAAAGCAAGTAG 749
20057 TGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 19998
750 CCGTGACCGGATTTAAGGACCAATGATGAGAGCTTTGAAGAGATCACTCCGCGTTA 809
19997 TGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 19938
810 TGTCTTGAAGCTACTTGGCTTACCGCTTGATGATTAACGTCGCGAAGAAAGCTAAATG 869
19937 TGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 19878
870 TTAAAGCGGTGCGGAATTTGTGCTGCTGAGAGAGTGAAGATCAGCTCTTGT 929
19877 TGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 19818
930 TGGGGGTTGACCCCTGAGAAATTATGAGAGAGCGTTTACAAAGACAGCTGCTGA 989
19817 TGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 19758
990 GCAAGGTGATCTTTTGTAGTACCCCAAGCAATTCAGAGAGTCAATTTGAAGTTA 1049
19757 TGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 19658
1050 CGAAGTTCACCTGCTCTTGTGCTCAAGCTTTATGAGTGAAGACCAAGCTTTTACA 1109
19697 TGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 19638
1110 GGATGCTGAT 1119
19637 TGTCTGCTGCT 19628

RESULT 15
AC115666/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
AC115666 200412 bp DNA linear HTG 15-NOV-2002
Rattus norvegicus clone CH230-261G15, WORKING DRAFT SEQUENCE, 4
unordered pieces.
AC115666
AC115666.5 GI:25012557
HTG; HTGS PHASE1; HTGS DRAFT; HTGS FULLTOP.
Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Muridae; Muridae; Muridae; Rattus.
1 (bases 1 to 200412)
Muzny D, Marie, Metzker, M. Lee, Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsdorfs, S., Amin, A., Angiano, D.,
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Cesar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Gebregorgis, B., Geer, K., Gill, R., Grady, M., Guerra, W., Guera, W.,
Guarinate, P., Haland, W., Hamill, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M.,
Hollins, B., Howells, S., Huily, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kovis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorenz, H., Lounsbury, H., Lozano, R. J., Lu, X., Ma, J.,
Maheshwari, M., Mahindaratne, M., Mahmood, M., Malloy, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
Mawhiney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
Nwokedi, O., Okwum, G., Olariu, A., Pal, S., Parks, K.,
Pastermak, S., Paul, H., Perez, A., Perez, L., Pflanz, C.,
Plopper, F., Poldinger, A., Popovic, D., Primus, E., Pu, L., L.,
Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R.,
Riley, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J.,
Sanders, W., Savary, G., Scherer, S., Scott, G., Shatman, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smales, D.,
Sneed, A., Sodergren, E., Song, X. Z., Sorelle, R., Sosa, J.,
Steinle, M., Strong, R., Sutton, A., Stratek, A., Tabors, Z., Usmani, K.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
Valas, R., Vera, V., Villaseca, D., Waldron, L., Walker, B., Wang, J.,
Wang, Q., Wang, S., Warren, R., Warren, R., Wei, X., Whitley, F.,
Williams, G., Willson, R., Wleczky, R., Wood, H., Wylie, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Zhou, D., von
Niederhausen, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,
Weinstock, G. and Gibbs, R. A.
Direct Submission
Unpublished
2 (bases 1 to 200412)
Worley, K. C.
Direct Submission
Submitted (22-MAR-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 10, 2005, 08:09:29 ; Search time 1361.27 Seconds
(without alignments)
11779.592 Million cell updates/sec

Title: US-10-600-070B-1

Perfect score: 2406

Sequence: 1 atggaagctctgagtcacgt.....gctctgtctctgcataca 2406

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_21:*

- 1: geneseqn1980s:*
- 2: geneseqn1990s:*
- 3: geneseqn2000s:*
- 4: geneseqn2001as:*
- 5: geneseqn2001bs:*
- 6: geneseqn2002as:*
- 7: geneseqn2002bs:*
- 8: geneseqn2003as:*
- 9: geneseqn2003bs:*
- 10: geneseqn2003cs:*
- 11: geneseqn2003ds:*
- 12: geneseqn2004as:*
- 13: geneseqn2004bs:*
- 14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2406	100.0	2406	12	ADJ38129
2	2402.8	99.9	2406	12	ADJ38129 Arabidops
3	2399.6	99.7	2406	12	ADJ38135 Arabidops
4	2399.6	99.7	2406	12	ADJ38208 Arabidops
5	2399.6	99.7	2406	12	ADJ38210 Arabidops
6	1724	71.7	2679	13	ADT14901 plant CDN
7	1720.8	71.5	3667	12	ADJ38130 Arabidops
8	482	20.0	561	12	ADJ38136 Arabidops
9	481	20.0	561	12	ADJ38212 Arabidops
10	328.8	13.7	631	12	ADJ38206 Arabidops
11	278.8	11.6	631	12	ADJ38264 Arabidops
12	224	9.3	537	12	ADJ38223 Arabidops
13	222.4	9.2	552	13	ACN48855 Cotton pr
14	207	8.6	545	12	ADJ38255 Arabidops
15	197.2	8.2	491	12	ADJ38254 Arabidops
16	195.8	8.1	647	12	ADJ38265 Arabidops
17	193.8	8.1	460	12	ADJ38267 Arabidops
18	190	7.9	446	12	ADJ38231 Arabidops
19	186.6	7.8	652	12	ADJ38266 Arabidops

20	178.6	7.4	607	13	ACN48781	ACN48781 Cotton pr
21	175	7.3	439	12	ADP93143	ADP93143 Cotton ex
22	170.4	7.1	608	12	ADJ38217	ADJ38217 Arabidops
23	170.2	7.1	653	12	ADJ38258	ADJ38258 Arabidops
24	169.8	7.1	622	12	ADJ38226	ADJ38226 Arabidops
25	165.4	6.9	527	12	ADJ38214	ADJ38214 Arabidops
26	158.6	6.6	871	12	ADJ38232	ADJ38232 Arabidops
27	155.4	6.5	307	12	ADJ38218	ADJ38218 Arabidops
28	150.4	6.3	418	12	ADJ38224	ADJ38224 Arabidops
29	144.6	6.0	653	12	ADJ38228	ADJ38228 Arabidops
30	144.2	6.0	479	12	ADJ38230	ADJ38230 Arabidops
31	130.8	5.4	420	12	ADJ38256	ADJ38256 Arabidops
32	128.6	5.3	480	12	ADJ38225	ADJ38225 Arabidops
33	127.6	5.3	535	12	ADJ38229	ADJ38229 Arabidops
34	124.2	5.2	360	12	ADJ38259	ADJ38259 Arabidops
35	118.8	4.9	336	12	ADJ38222	ADJ38222 Arabidops
36	117.8	4.9	309	12	ADJ38221	ADJ38221 Arabidops
37	111.4	4.6	540	12	ADJ38233	ADJ38233 Arabidops
38	111.2	4.6	604	12	ADJ38227	ADJ38227 Arabidops
39	110.6	4.6	606	12	ADJ38257	ADJ38257 Arabidops
40	96.8	4.0	295	12	ADJ38213	ADJ38213 Arabidops
41	92.6	3.8	187	12	ADJ38216	ADJ38216 Arabidops
42	83.2	3.5	300	12	ADJ38260	ADJ38260 Arabidops
43	77.6	3.2	400	12	ADJ38263	ADJ38263 Arabidops
44	76.8	3.2	416	12	ADJ38219	ADJ38219 Arabidops
45	75.2	3.1	549	12	ADJ38261	ADJ38261 Arabidops

ALIGNMENTS

RESULT 1	ADJ38129	standard; cDNA; 2406 BP.
ID	ADJ38129;	
AC	ADJ38129;	
DT	06-MAY-2004	(first entry)
DE	Arabidopsis thaliana Arc6-1 cDNA SeqID1.	
KW	prokaryotic type; plastid division; Ftn2; ARCE; ARCS; Fzo; plant cell; agronomic; horticultural; crop plant; ornamental plant; woody plant; herbicide target; gene; ss.	
OS	Arabidopsis thaliana.	
XX	WO2004001003-A2.	
XX	31-DEC-2003.	
XX	20-JUN-2003; 2003WO-US019536.	
XX	20-JUN-2002; 2002US-0390140P.	
XX	09-AUG-2002; 2002US-0402242P.	
XX	20-JUN-2003; 2003US-00600070.	
XX	(UNMS) UNIV MICHIGAN STATE.	
XX	Osteryoung KW, Vittha S, Kokscharova OA, Gao H;	
XX	WPI; 2004-082486/08.	
XX	P-PSDB; ADJ38202.	
XX	New isolated Ftn2, ARCS and/or Fzo-like nucleic acid sequences, useful for further characterizing plastid division in plant cells, and in varying agronomic and horticultural characteristics of economically important plants.	
XX	Claim 1; SEQ ID NO 1; 287pp; English.	
XX	This invention relates to novel prokaryotic type or plastid division and related genes and proteins. In particular, the invention relates to novel	

QY 1981 GGGTCAGTCAGAGCTGACGATTCCAGAACCTTCCAGAAATGATGCTAGAGCTGCAAG 2040
DB 1981 GGGTCAGTCAGAGCTGACGATTCCAGAACCTTCCAGAAATGATGCTAGAGCTGCAAG 2040
QY 2041 AATATAGTATCCAGTGGCAGAAAGTTAAGTCTGCGCTTTGGCCCTGATCCAGCCGATA 2100
DB 2041 AATATAGTATCCAGTGGCAGAAAGTTAAGTCTGCGCTTTGGCCCTGATCCAGCCGATA 2100
QY 2101 GAAATGTTACGAGAGTGGTGGATGGGCGAAATGCTGAAGATTGGACTGACAGACAGACT 2160
DB 2101 GAAATGTTACGAGAGTGGTGGATGGGCGAAATGCTGAAGATTGGACTGACAGACAGACT 2160
QY 2161 GAACTGCGCAGCTGGGCTGGTTATATATATACACTGTTGAAACTATCTGTTGACAGT 2220
DB 2161 GAACTGCGCAGCTGGGCTGGTTATATATATACACTGTTGAAACTATCTGTTGACAGT 2220
QY 2221 GTGACAGTCTCAGCAGATGGAACCGGTGCTGCTGGTGAAGCAACTCTGAGAGAGTCTGCT 2280
DB 2221 GTGACAGTCTCAGCAGATGGAACCGGTGCTGCTGGTGAAGCAACTCTGAGAGAGTCTGCT 2280
QY 2281 TGTCTATCTGATTTGGTTTCATCCAGAAAACAATGCTAGTGTGAGAACTACACAACA 2340
DB 2281 TGTCTATCTGATTTGGTTTCATCCAGAAAACAATGCTAGTGTGAGAACTACACAACA 2340
QY 2341 AGATACGAAGTTTCTGCTCCAGTCCAGGCTGGAAGAAATCACTGAAGGCTCTGTTCTTCA 2400
DB 2341 AGATACGAAGTTTCTGCTCCAGTCCAGGCTGGAAGAAATCACTGAAGGCTCTGTTCTTCA 2400
QY 2401 TCATPAA 2406
DB 2401 TCATPAA 2406

RESULT 2
ID ADJ38135 standard; cDNA; 2406 BP.
XX ADJ38135;
AC

DT 06-MAY-2004 (first entry)

XX Arabidopsis thaliana AetFn2 cDNA SeqID9.

XX prokaryotic type; plastid division; Ftn2; ARC6; ARC5; Fzo; plant cell;
KW agronomic; horticultural; crop plant; ornamental plant; woody plant;
KW herbicide target; gene; ss.

OS Arabidopsis thaliana.

XX WO2004001003-A2.

XX 31-DEC-2003.

XX 20-JUN-2003; 2003WO-US019536.

XX 20-JUN-2002; 2002US-0390140P.

XX 09-AUG-2002; 2002US-0402242P.

XX 20-JUN-2003; 2003US-00600070.

XX (UNMS) UNIV MICHIGAN STATE.

XX Oseeryoung KW, Vittha S, Koksharova OA, Gao H;

XX WPI; 2004-082486/08.

XX P-PSDB; ADJ38203.

XX New isolated Ftn2, ARC5 and/or Fzo-like nucleic acid sequences, useful
PT for further characterizing plastid division in plant cells, and in
PT varying agronomic and horticultural characteristics of economically
PT important plants.

XX Claim 1; SEQ ID NO 9; 287bp; English.

CC This invention relates to novel prokaryotic type or plastid division and
CC related genes and proteins. In particular, the invention relates to novel
CC Ftn2 (ARC6), ARC5 and Fzo-like genes and polypeptides. The methods and
CC compositions of the present invention are useful for further
CC characterizing plastid division in plant cells, in order to vary
CC agronomic and horticultural characteristics of economically important
CC plants, such as crop, ornamental and woody plants. They can also be used
CC as herbicide targets. The present sequence is a cDNA sequence which is
CC related to the invention.

XX Sequence 2406 BP; 611 A; 491 C; 619 G; 685 T; 0 U; 0 Other;

Query Match 99.9%; Score 2402.8; DB 12; Length 2406;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2404; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGAAGTCTGATCACTGCTGCGATGCTCTCCCATTCCTCAATTATGCGGATTCACA 60
DB 1 ATGGAAGTCTGATCACTGCTGCGATGCTCTCCCATTCCTCAATTATGCGGATTCACA 60
QY 61 CCGGCGAGCAAAAGCTCCGACGTAGCCACAACACTCTCAACTATCTGCTCCGACAGC 120
DB 61 CCGGCGAGCAAAAGCTCCGACGTAGCCACAACACTCTCAACTATCTGCTCCGACAGC 120
QY 121 AATAGGCGCAGCCGCTTCTCTCCGACTCAATTTCACCTCCGATTCCTCTCTCTCC 180
DB 121 AATAGGCGCAGCCGCTTCTCTCCGACTCAATTTCACCTCCGATTCCTCTCTCTCC 180
QY 181 TTGCGCACCCGACACACACACCGCACTCTGCTCTCTGCGACCATATATGATGCTGCC 240
DB 181 TTGCGCACCCGACACACACCGCACTCTGCTCTCTGCGACCATATATGATGCTGCC 240
QY 181 TTGCGCACCCGACACACACCGCACTCTGCTCTCTGCGACCATATATGATGCTGCC 240
DB 181 TTGCGCACCCGACACACACCGCACTCTGCTCTCTGCGACCATATATGATGCTGCC 240
QY 241 GAACGCCACGTCCTCCATCCCATTTGATTTTACACAGGATATAGAGCTCAACAATTC 300
DB 241 GAACGCCACGTCCTCCATCCCATTTGATTTTACACAGGATATAGAGCTCAACAATTC 300
QY 301 TTAACCGATGGAATCAGAAAGCATTCGAAGCTTGGTGAACCCGCGCAATTCGGT 360
DB 301 TTAACCGATGGAATCAGAAAGCATTCGAAGCTTGGTGAACCCGCGCAATTCGGT 360
QY 361 TTGACCGACGACGCTTTAATATAGCCGAGACAGATTCTTCAAGTGTCTTGGAAACTGTG 420
DB 361 TTGACCGACGACGCTTTAATATAGCCGAGACAGATTCTTCAAGTGTCTTGGAAACTGTG 420
QY 421 TCTAATCCTGCGTCTAGAAAGAGTATCAATGAAGTCTTCTTGAATGAAGAACTACA 480
DB 421 TCTAATCCTGCGTCTAGAAAGAGTATCAATGAAGTCTTCTTGAATGAAGAACTACA 480
QY 481 GTCATCACTGATGTTCTTGGGATGAAGGTTCTGGGCTCTCTGTGATTCGAAGAGGT 540
DB 481 GTCATCACTGATGTTCTTGGGATGAAGGTTCTGGGCTCTCTGTGATTCGAAGAGGT 540
QY 541 GGTGAGCTGAGATGATGTTCTTGGGATGAAGGTTCTGGGCTCTCTGTGATTCGAAGAGGT 600
DB 541 GGTGAGCTGAGATGATGTTCTTGGGATGAAGGTTCTGGGCTCTCTGTGATTCGAAGAGGT 600
QY 601 TCGTTTAAAGCAAGATGCTTTAGTATGAGGCTCTGCTTAAAGAGAGGTTCTTAAAG 660
DB 601 TCGTTTAAAGCAAGATGCTTTAGTATGAGGCTCTGCTTAAAGAGAGGTTCTTAAAG 660
QY 661 GCTATGCGATTGATTCACCTGATTTTATTAATGCTTATGAGTTTGTGAGAAAGCTTTG 720
DB 661 GCTATGCGATTGATTCACCTGATTTTATTAATGCTTATGAGTTTGTGAGAAAGCTTTG 720
QY 721 AAGCTTTTAAAGAT 780
DB 721 AAGCTTTTAAAGAT 780
QY 781 GAGACTTTGGAAGAGATCACTCCGCTTATGCTTGAAGCTACTTGGCTTACCGCTTGGT 840
DB 781 GAGACTTTGGAAGAGATCACTCCGCTTATGCTTGAAGCTACTTGGCTTACCGCTTGGT 840
QY 841 GATGATTAACGCTCGGAAAGACATAAATGTTTAAAGCGGTGTGCGGAATATTTTGTGCT 900

Db 841 GATGATTCGCTGCGCAAAAAGCTAAATGATTTAAACGGGTGCGGAAATTTTGTGCTCT 900
Qy GTTGAGAGAGGAGGAGCATCAGCTCTGTGTGGGGTTTGACCCGGAGAAAGTTATGAT 960
Db GTTGAGAGAGGAGGAGCATCAGCTCTGTGTGGGGTTTGACCCGGAGAAAGTTATGAT 960
Qy GAGGCGTTTTTACGAATGACAGCTGCTGAGCAGTTGATCTTTTGTAGTACCCCAAGC 1020
Db GAGGCGTTTTTATGAATGACAGCTGCTGAGCAGTTGATCTTTTGTAGTACCCCAAGC 1020
Qy AATATTTCCAGACAGATCTTTTGAAGTTTACGAAGTTGACCTTGTGCTGCTCAAGCT 1080
Db AATATTTCCAGACAGATCTTTTGAAGTTTACGAAGTTGACCTTGTGCTGCTCAAGCT 1080
Qy TTTATTTGTAAGAACCCACACTTTTACAGATGCTGATTAAGCAATTCAGCACTTCAAG 1140
Db TTTATTTGTAAGAACCCACACTTTTACAGATGCTGATTAAGCAATTCAGCACTTCAAG 1140
Qy CAGGCTAAGGTAATGCTATGAGATTCCTGCGATGTTGATGATACACGGAATATTGG 1200
Db CAGGCTAAGGTAATGCTATGAGATTCCTGCGATGTTGATGATACACGGAATATTGG 1200
Qy GAGATTAAGCTTCGCTCTGAGAAAGGGGACTCTGTGACATGCTTATAGCAAAAGTTGAT 1260
Db GAGATTAAGCTTCGCTCTGAGAAAGGGGACTCTGTGACATGCTTATAGCAAAAGTTGAT 1260
Qy TGCCGATGCTGCTGCTGCTGAGACAGTGAAGTTCAATATTAAGCAATTCAGCACTT 1320
Db TGCCGATGCTGCTGCTGCTGAGACAGTGAAGTTCAATATTAAGCAATTCAGCACTT 1320
Qy GAGTTTGTGTTTGAAGAAATTCGAATGCTGATGATGATGATGATGATGATGATGAT 1380
Db GAGTTTGTGTTTGAAGAAATTCGAATGCTGATGATGATGATGATGATGATGATGAT 1380
Qy TTTGTTGAGAAAGCTGCTGCTGAGGGGTTCTTTCTAGCTTCAAGACACCAAAAGAT 1440
Db TTTGTTGAGAAAGCTGCTGCTGAGGGGTTCTTTCTAGCTTCAAGACACCAAAAGAT 1440
Qy AAATTTTAACTCGGGGAACTACTATGATGATGATGATGATGATGATGATGATGAT 1500
Db AAATTTTAACTCGGGGAACTACTATGATGATGATGATGATGATGATGATGATGAT 1500
Qy GAGGTAAGTTCAGGGTTCTCTTTAGCTGCTGCTGCACTATGAGCAAGATTGAGCCGAG 1560
Db GAGGTAAGTTCAGGGTTCTCTTTAGCTGCTGCTGCACTATGAGCAAGATTGAGCCGAG 1560
Qy CATGTGAAAGCTAGTGTGCTATGCAAGCACTGCAAGAAAGTTTCTTCCGCTATACAGAT 1620
Db CATGTGAAAGCTAGTGTGCTATGCAAGCACTGCAAGAAAGTTTCTTCCGCTATACAGAT 1620
Qy AGAAACTCGGCTGAAACCCAGATGTCAGAGACAGTGTATGATGATGATGATGATGAT 1680
Db AGAAACTCGGCTGAAACCCAGATGTCAGAGACAGTGTATGATGATGATGATGATGAT 1680
Qy AACATATGAGGCGCTGATGCTGAGCTGCTGCTTATTTATGAGAAGCTGTAAGCCCTCT 1740
Db AACATATGAGGCGCTGATGCTGAGCTGCTGCTTATTTATGAGAAGCTGTAAGCCCTCT 1740
Qy GAAAACCTTTGAACCTAATGATTAATGCAATGAGCTGGGGTCTCAAGAGTAGGTTGAT 1800
Db GAAAACCTTTGAACCTAATGATTAATGCAATGAGCTGGGGTCTCAAGAGTAGGTTGAT 1800
Qy GAAACTACTGTTGAAATGCTCGTGTGCTGATGATGATGATGATGATGATGATGAT 1860
Db GAAACTACTGTTGAAATGCTCGTGTGCTGATGATGATGATGATGATGATGATGAT 1860
Qy GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920
Db GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920
Qy AGCTCATCTTTTCAACGCAAGGATGATGATGATGATGATGATGATGATGATGAT 1980
Db AGCTCATCTTTTCAACGCAAGGATGATGATGATGATGATGATGATGATGATGAT 1980

Db 1921 AGCTCATCTTTTCAACGCAAGGATGATGATGATGATGATGATGATGATGAT 1980
Qy GGGTCAGTCAGAGCTGACGATTCAGAAAGCACTTCCAGAAATGATGCTAGGACTGACAG 2040
Db GGGTCAGTCAGAGCTGACGATTCAGAAAGCACTTCCAGAAATGATGCTAGGACTGACAG 2040
Qy AATATAGTATCCAGAGTGGCAGAAAGATTAACTCTGCTGCTTTTGGGCTGATCACCGCAT 2100
Db AATATAGTATCCAGAGTGGCAGAAAGATTAACTCTGCTGCTTTTGGGCTGATCACCGCAT 2100
Qy GAAATGTTTACCAAGAGTTTGGATGGCGGAATGCTGAAGATTGAGTACAGACAGCT 2160
Db GAAATGTTTACCAAGAGTTTGGATGGCGGAATGCTGAAGATTGAGTACAGACAGCT 2160
Qy GAACTGCGCAGCTTGGGTTGTTTATGATATACATGCTGTTGAAACTATCTGTTGACAGT 2220
Db GAACTGCGCAGCTTGGGTTGTTTATGATATACATGCTGTTGAAACTATCTGTTGACAGT 2220
Qy GTGACAGTCTCAGAGATGAAACCCGCTCTGCTGAGAGCACTGAGAGAGTCTGCT 2280
Db GTGACAGTCTCAGAGATGAAACCCGCTCTGCTGAGAGCACTGAGAGAGTCTGCT 2280
Qy TGTCTATCTGATTTGCTTCAATCCAGAAACAAATGCTACTGATGTCAGAACTTACACACA 2340
Db TGTCTATCTGATTTGCTTCAATCCAGAAACAAATGCTACTGATGTCAGAACTTACACACA 2340
Qy AGATACGAAGTTTCTGCTCCAGTCAAGGCTGAGAAATCACTGAAGGCTCTGTTCTTGA 2400
Db AGATACGAAGTTTCTGCTCCAGTCAAGGCTGAGAAATCACTGAAGGCTCTGTTCTTGA 2400
Qy 2401 TCATPA 2406
Db 2401 TCATPA 2406

RESULT 3
ADJ38208
ID ADJ38208 standard; DNA; 2406 BP.
XX
XX AC ADJ38208;
XX
XX 06-MAY-2004 (first entry)
XX
XX
XX Plastid division-related *Arc6* orthologue gene 2.
XX
XX prokaryotic type; plastid division; *Ftn2*; *ARC6*; *ARC5*; *Fzo*; plant cell;
XX agronomic; horticultural; crop plant; ornamental plant; woody plant;
XX herbicide target; gene; ds.
XX
XX *Arabidopsis thaliana*.
XX
XX OS
XX PN
XX PN
XX PD
XX PD 31-DEC-2003.
XX
XX 20-JUN-2003; 2003WO-US019536.
XX PF
XX 20-JUN-2002; 2002US-0390140P.
XX PR 09-AUG-2002; 2002US-0402242P.
XX PR 20-JUN-2003; 2003US-0060070.
XX
XX (UNMS) UNIV MICHIGAN STATE.
XX PA
XX
XX OeserYoung KW, Vicha S, Koksharova OA, Gao H;
XX
XX WPI; 2004-082486/08.
XX DR
XX P-PSDB; ADJ38207.
XX
XX New isolated *Ftn2*, *ARC5* and/or *Fzo*-like nucleic acid sequences, useful
XX PT for further characterizing plastid division in plant cells, and in
XX PT varying agronomic and horticultural characteristics of economically
XX PT important plants.
XX

QY 1921 AGCTCATCTTTTCAACGCAAGATATGTTTCTTATGATCTGATGTCGTAACATA 1980
DB 1921 AGCTCATCTTTTCAACGCAAGATATGTTTCTTATGATCTGATGTCGTAACATA 1980
QY 1981 GGGTCAGTCAGAGCTGAGAGATTCGAAAGCACTTCCAGAAATGATGCTAGACCTGCAGAG 2040
DB 1981 GGGTCAGTCAGAGCTGAGAGATTCGAAAGCACTTCCAGAAATGATGCTAGACCTGCAGAG 2040
QY 2041 AATATAGATCCAGAGTGGCAGAAATTAAGTCTGCTTTGGGCTGATCAGCGATA 2100
DB 2041 AATATAGATCCAGAGTGGCAGAAATTAAGTCTGCTTTGGGCTGATCAGCGATA 2100
QY 2101 GAAATGTTACCAAGCTTTTGAATGGCGCAATGCTGAAATTTGACTGACAGACAGCT 2160
DB 2101 GAAATGTTACCAAGCTTTTGAATGGCGCAATGCTGAAATTTGACTGACAGACAGCT 2160
QY 2161 GAACTGGCGAGCTGGTGGTGTATGATATACCTGTTGAACATCTGTTGACAGT 2220
DB 2161 GAACTGGCGAGCTGGTGGTGTATGATATACCTGTTGAACATCTGTTGACAGT 2220
QY 2221 GTGACAGTCTCAGAGATGAAACCGTGTCTGCTGAGGAGCACTCTGAGAGAGTGTCT 2280
DB 2221 GTGACAGTCTCAGAGATGAAACCGTGTCTCTGCTGAGGAGCACTCTGAGAGAGTGTCT 2280
QY 2281 TGCTATCTGATTTGGTTTCAATCCAGAAACAACTGCTGATGTCAGAACTTACACAA 2340
DB 2281 TGCTATCTGATTTGGTTTCAATCCAGAAACAACTGCTGATGTCAGAACTTACACAA 2340
QY 2341 AGATACGAGATTTCTGGTCCAGTCCAGGCTGAGAAATCACTGAAAGCTCTGTTCTTGA 2400
DB 2341 AGATACGAGATTTCTGGTCCAGTCCAGGCTGAGAAATCACTGAAAGCTCTGTTCTTGA 2400
QY 2401 TCATTA 2406
DB 2401 TCATTA 2406

RESULT 4
ADJ38210
ID ADJ38210 standard; DNA; 2637 BP.

AC ADJ38210;
DT 06-MAY-2004 (first entry)
XX Plastid division-related Arc6 orthologue gene 3.
XX prokaryotic type; plastid division; Ftn2; ARC6; ARC5; Fzo; plant cell;
KM agronomic; horticultural; crop plant; ornamental plant; woody plant;
KM herbicide target; gene; ds.
XX Arabidopsis thaliana.
OS
XX
PN WO2004001003-A2.
PD 31-DEC-2003.
PF 20-JUN-2003; 2003WO-US019536.
XX 20-JUN-2002; 2002US-0390140P.
PR 09-AUG-2002; 2002US-0402242P.
PR 20-JUN-2003; 2003US-00600070.
XX
PA (UNMS) UNIV MICHIGAN STATE.
PI Oseeryoung KW, Vitha S, Kokeharova OA, Gao H;
DR WPI; 2004-082486/08.
DR P-PSDB; ADJ38209.
XX
PT New isolated Ftn2, ARCS and/or Fzo-like nucleic acid sequences, useful
PT for further characterizing plastid division in plant cells, and in
PT varying agronomic and horticultural characteristics of economically

PT important plants.
XX Disclosure; Fig 8; 287pp; English.
PS
XX
CC This invention relates to novel prokaryotic type or plastid division and
CC related genes and proteins. In particular, the invention relates to novel
CC Ftn2 (ARC6), ARCS and Fzo-like genes and polypeptides. The methods and
CC compositions of the present invention are useful for further
CC characterizing plastid division in plant cells, in order to vary
CC agronomic and horticultural characteristics of economically important
CC plants, such as crop, ornamental and woody plants. They can also be used
CC as herbicide targets. The present sequence is that of a gene which is
CC related to the invention.
XX
SQ Sequence 2637 BP; 706 A; 535 C; 644 G; 752 T; 0 U; 0 Other;
Query Match 99.7%; Score 2399.6; DB 12; Length 2637;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2402; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 ATGGAAGCTCTGAGTCACGTCGGCATTGGTCTCTCCCATTCGAATTATGCCATTACCA 60
DB 114 ATGGAAGCTCTGAGTCACGTCGGCATTGGTCTCTCCCATTCGAATTATGCCATTACCA 173
QY 61 CCGGCGAGCAAAAGCTCCGACGTAGCCACAACAACCTTACAACTATCTGCTCCGACG 120
DB 174 CCGGCGAGCAAAAGCTCCGACGTAGCCACAACAACCTTACAACTATCTGCTCCGACG 233
QY 121 AATAGGCGGACCGCTCTCTCTCCGACTTGAATTTCAACCTCGATTCTCTCTCTCC 180
DB 234 AATAGGCGGACCGCTCTCTCTCCGACTTGAATTTCAACCTCGATTCTCTCTCTCC 293
QY 181 TTGCGCACCGCCACACACACCGCACCTCTGCTCTGCGACCATATTTGATTCGCTCC 240
DB 294 TTGCGCACCGCCACACACACCGCACCTCTGCTCTGCGACCATATTTGATTCGCTCC 353
QY 241 GAAAGCCACGTCCTCCATCCCATTTGATTTTACACAGATTTAGAGAGCTCAACACATTT 300
DB 354 GAAAGCCACGTCCTCCATCCCATTTGATTTTACACAGATTTAGAGAGCTCAACACATTT 413
QY 301 TTAAACGATGGAATCCAGAAAGCATTCGAGCTTTCGAAACCGCGCAATTCGCT 360
DB 414 TTAAACGATGGAATCCAGAAAGCATTCGAGCTTTCGAAACCGCGCAATTCGCT 473
QY 361 TTACGACGACGCTTTAATCAACCGGAGACAGATTTCTTCAAGCTGCTCGAAACTCTG 420
DB 474 TTACGACGACGCTTTAATCAACCGGAGACAGATTTCTTCAAGCTGCTCGAAACTCTG 533
QY 421 TCTAATCTCGCTGAGAAAGATACAGATGAAGCTTCTTGATGATGAAGAGCTTACA 480
DB 534 TCTAATCTCGCTGAGAAAGATACAGATGAAGCTTCTTGATGATGAAGAGCTTACA 593
QY 481 GTGATCATGATGTCCTTGGGATTAAGTTCCTGGGCTCTGCTGATATGCAAGAGGT 540
DB 594 GTGATCATGATGTCCTTGGGATTAAGTTCCTGGGCTCTGCTGATATGCAAGAGGT 653
QY 541 GGTGAGACTGAGATAGTCTTTCGGGTTGTGAGGCTCTGCTTAAAGAGAGGTGCTTAAG 600
DB 654 GGTGAGACTGAGATAGTCTTTCGGGTTGTGAGGCTCTGCTTAAAGAGAGGTGCTTAAG 713
QY 601 TCGTTTAAAGCAAGATGCTTTTAAATTAAGCGCTTGTCTCGATGCTCGAGGAGAT 660
DB 714 TCGTTTAAAGCAAGATGCTTTTAAATTAAGCGCTTGTCTCGATGCTCGAGGAGAT 773
QY 661 GCTATGCAATTTGATCCACCTGATTTTAACTAGCTTATAGTTTGTGGAAGCTTGT 720
DB 774 GCTATGCAATTTGATCCACCTGATTTTAACTAGCTTATAGTTTGTGGAAGCTTGT 833
QY 721 AAGCTTTTAAAGAGAGAGAGCAAGTATGCTTGCACCGGATTTACGTGACCAATTTGAT 780
DB 834 AAGCTTTTAAAGAGAGAGAGCAAGTATGCTTGCACCGGATTTACGTGACCAATTTGAT 893
QY 781 GAGACTTTGAGAGAGATCACTCCGCGTTATGCTTGGAGCTACTTGGCTTACCGCTTGGT 840

Db 894 GAGACTTTGGAAGAGATCACTCGCGTTATGTCCTTGAGACTACTGGCTTACCGCTTGGCT 953
Qy 841 GATGATTTACGTCGCGAAAAGACTAAATGGTTTAAAGCGGTGTGCGAAATATTTTGTGTCT 900
Db 954 GATGATTTACGTCGCGAAAAGACTAAATGGTTTAAAGCGGTGTGCGAAATATTTTGTGTCT 1013
Qy 901 GTTGGAGAGGTGAGCATCACTCTTGTGGGGGTTTGACCCGTGAGAGATTATGAAT 960
Db 1014 GTTGGAGAGGTGAGCATCACTCTTGTGGGGGTTTGACCCGTGAGAGATTATGAAT 1073
Qy 961 GAGGGGTTTTTACGATGACAGTGTGAGAGGTGATCTTTTGTACTACCCCAAC 1020
Db 1074 GAGGGGTTTTTACGATGACAGTGTGAGAGGTGATCTTTTGTACTACCCCAAC 1133
Qy 1021 AATATTTCCAGAGAGCTCACTTGAAGTTTACGAAGTTGACCTTGTGGCTCAAGCT 1080
Db 1134 AATATTTCCAGAGAGCTCACTTGAAGTTTACGAAGTTGACCTTGTGGCTCAAGCT 1193
Qy 1081 TTTATTTGTTAGAGAGCCACCTTTTACAGATGCTGATAGCAATTTCCAGCACTTCAG 1140
Db 1194 TTTATTTGTTAGAGAGCCACCTTTTACAGATGCTGATAGCAATTTCCAGCACTTCAG 1253
Qy 1141 CAGGCTAAGGTAATGCTATGAGAGATCTCGGAGTGTGTATGATACAGCAATTAATGG 1200
Db 1254 CAGGCTAAGGTAATGCTATGAGAGATCTCGGAGTGTGTATGATACAGCAATTAATGG 1313
Qy 1201 GAGATAGACTTCGCTTAGAAAAGGGGACTCTGAGCACTGCTTATAGGCAAGTTGATGA 1260
Db 1314 GAGATAGACTTCGCTTAGAAAAGGGGACTCTGAGCACTGCTTATAGGCAAGTTGATGA 1373
Qy 1261 TCGCGTAATGTTGGGCTTAGACAGTAGAGATTCAATATAGAAATCCAGCTATTTGTG 1320
Db 1374 TCGCGTAATGTTGGGCTTAGACAGTAGAGATTCAATATAGAAATCCAGCTATTTGTG 1433
Qy 1321 GAGTTTGTTTTGGAGAAATTCAAATGTGTATGCAATGATGATCTCCCTGGAATATGCAA 1380
Db 1434 GAGTTTGTTTTGGAGAAATTCAAATGTGTATGCAATGATGATCTCCCTGGAATATGCAA 1493
Qy 1381 TTGTTGAAAACCTGTGTGAGGGGTTGTTCTTCTAGGTTGAGAGACCAAGATTA 1440
Db 1494 TTGTTGAAAACCTGTGTGAGGGGTTGTTCTTCTAGGTTGAGAGACCAAGATTA 1553
Qy 1441 AAATTTAACTCGGGGACTACTATGATGATCCTATGTTTGTAGTACTTGGAAAAGTG 1500
Db 1554 AAATTTAACTCGGGGACTACTATGATGATCCTATGTTTGTAGTACTTGGAAAAGTG 1613
Qy 1501 GAGGTAGTTCAAGGTTCTTCTTACCTGTGCTGCACTATGCAAGATTGAGCCGAG 1560
Db 1614 GAGGTAGTTCAAGGTTCTTCTTACCTGTGCTGCACTATGCAAGATTGAGCCGAG 1673
Qy 1561 CATGTGAAGCTAGTGTCTATGCAAGGCACTGCAAGAAAGTTTCTTCCGCTATACAGAT 1620
Db 1674 CATGTGAAGCTAGTGTCTATGCAAGGCACTGCAAGAAAGTTTCTTCCGCTATACAGAT 1733
Qy 1621 AGAAACTCGGCTGAACCCAGAGATGTGCAAGAGAGTGTATGATGATCTTGTGT 1680
Db 1734 AGAAACTCGGCTGAACCCAGAGATGTGCAAGAGAGTGTATGATGATCTTGTGT 1793
Qy 1681 AAACAATGAGCCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
Db 1794 AAACAATGAGCCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1853
Qy 1741 GAAAACCTTTGAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800
Db 1854 GAAAACCTTTGAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1913
Qy 1801 GAAACTACTGTTGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
Db 1914 GAAACTACTGTTGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1973
Qy 1861 GCTGCTGTGTGCAATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1920

Db 1974 GCTGCTGTGTGCAATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2033
Qy 1921 AGCTCATCTTTTCAACGCAAGGATATGATGATGATGATGATGATGATGATGATGATGATGAT 1980
Db 2034 AGCTCATCTTTTCAACGCAAGGATATGATGATGATGATGATGATGATGATGATGATGATGAT 2093
Qy 1981 GGGTCAGTCAGAGCTGACGATTCAGAAAGCACTTCCAGAAATGATGATGATGATGATGATGAT 2040
Db 2094 GGGTCAGTCAGAGCTGACGATTCAGAAAGCACTTCCAGAAATGATGATGATGATGATGATGATGAT 2153
Qy 2041 AATATGATTCGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2100
Db 2154 AATATGATTCGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2213
Qy 2101 GAAATGTTTACAGAGGTTTGTGATGGGCGAATGCTGAGATTTGAGCTGACAGAGAGCT 2160
Db 2214 GAAATGTTTACAGAGGTTTGTGATGGGCGAATGCTGAGATTTGAGCTGACAGAGAGCT 2273
Qy 2161 GAACTGCGCAGCTTGGGTTGTTATGATTAATTAACACTGTTGAAACTATCTGTTGACAGT 2220
Db 2274 GAACTGCGCAGCTTGGGTTGTTATGATTAATTAACACTGTTGAAACTATCTGTTGACAGT 2333
Qy 2221 GTGACAGTCTCAGAGATGAAACCCGTGCTGTGTGAGAGCACTCTGAGAGATGCTGT 2280
Db 2334 GTGACAGTCTCAGAGATGAAACCCGTGCTGTGTGAGAGCACTCTGAGAGATGCTGT 2393
Qy 2281 TGTCTATCTGATTTGGTTCAATCAGAAAACAATGCTATGATGATGATGATGATGATGATGATGAT 2340
Db 2394 TGTCTATCTGATTTGGTTCAATCAGAAAACAATGCTATGATGATGATGATGATGATGATGATGAT 2453
Qy 2341 AGATACGAAGTTTTCGTGCTCAAGTCAAGGTTGAAAATCACTGAAGGCTGTGTTTGTGCA 2400
Db 2454 AGATACGAAGTTTTCGTGCTCAAGTCAAGGTTGAAAATCACTGAAGGCTGTGTTTGTGCA 2513
Qy 2401 TCATTA 2406
Db 2514 TCATTA 2519

RESULT 5
ADT14901
ID ADT14901 standard; cDNA, 2679 BP.
XX
AC ADT14901;
XX
DT 13-JAN-2005 (first entry)
XX
DE Plant cDNA, Seq ID 227.
XX
KW Plant; ss; gene; transgenic; cold tolerance; growth rate;
KW drought tolerance; disease resistance; galactomannan production;
KW plant growth regulator; heat tolerance; herbicide tolerance;
KW lignin production; extreme osmotic condition tolerance;
KW pathogens resistance; pest resistance; yield improvement; seed oil yield;
KW seed protein yield.
OS
OS Viridiplantae.
XX
PN US2004216190-A1.
XX
PD 28-OCT-2004.
XX
PF 18-DEC-2003; 2003US-00739930.
XX
PR 28-APR-2003; 2003US-00424599.
PR 28-APR-2003; 2003US-00425115.
XX
PA (KOVA/) KOVALIC D K.
XX
XX Kovalic DK;
PI
XX
DR WPI; 2004-757369/74.
XX

PT New recombinant DNA constructs useful in the field of biochemistry and
PT genetics, and in particular for producing transgenic plants with improved
PT biological characteristics.

PS Claim 1; SEQ ID NO 227; 14bp; English.

XX The invention relates a recombinant DNA construct comprising a
CC polynucleotide having any of 5544 nucleotide sequences (CDNAs SEQ ID NO:
CC 1-5544) and encoding a polypeptide with any of 5544 amino acid sequences
CC (SEQ ID NO: 5545-11088). The CDNAs and proteins are from corn, soybean,
CC Arabidopsis, wheat and rape but the specification does not indicate which
CC sequences is derived from which organism. Also included is a method of
CC producing a plant having an improved property, comprising transforming a
CC plant with a recombinant DNA construct comprising a promoter region
CC functional in a plant cell operably joined to a polynucleotide encoding a
CC polypeptide associated with the property, and growing the transformed
CC plant. The property is selected from improving plant cold tolerance, for
CC manipulating growth rate in plant cells by modification of the cell cycle
CC pathway, for improving plant drought tolerance, for providing increased
CC resistance to plant disease, for galactomanan production, for production
CC of plant growth regulators, for improving plant heat tolerance, for
CC improving plant tolerance to herbicides, for increasing the rate of
CC homologous recombination in plants, for lignin production, for improving
CC plant tolerance to extreme osmotic conditions, for improving plant
CC tolerance to pathogens or pests, for yield improvement by modification of
CC photosynthesis, for modifying seed oil yield and/or content, for
CC modifying seed protein yield and/or content, for yield improvement by
CC modification of carbohydrate, nitrogen or phosphorus use and/or uptake
CC and for yield improvement by providing improved plant growth and
CC development under at least one stress condition. The polynucleotide may
CC also encode a plant transcription factor. The methods and compositions of
CC the present invention are useful in the field of biochemistry and
CC genetics, in particular for producing transgenic plants with improved
CC biological characteristics such as increased yield, improved nitrogen
CC flow, increasing plant tolerance to cold or heat, improving plant
CC tolerance to extreme osmotic and drought conditions, and improving plant
CC tolerance to plant pests or pathogens. They can also be used in physical
CC arrays of molecules, plant breeding markers, computer-based storage and
CC analysis systems. The present sequence is one of the 5544 plant cDNA
CC sequences of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?docid=20040216190.

XX Sequence 2679 BP; 712 A; 545 C; 653 G; 769 T; 0 U; 0 Other;

Query Match 99.7%; Score 2399.6; DB 13; Length 2679;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 2402; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGGAAGCTCTGAGTCAAGTGGGATTGGTCTCTCCCATTCATTAATGCGATTACCA 60
DB 115 ATGGAAGCTCTGAGTCAAGTGGGATTGGTCTCTCCCATTCATTAATGCGATTACCA 174
QY 61 CCGGCGACGACAAAGCTCCGACGTAAGCACAACCTCTACACTATCTGTCGGCAGC 120
DB 175 CCGGCGACGACAAAGCTCCGACGTAAGCACAACCTCTACACTATCTGTCGGCAGC 234
QY 121 AAATGGGCGGACGCTCTCTCTCGACTTCAATTCACTCCGATTCCTCTCCCTCC 180
DB 235 AAATGGGCGGACGCTCTCTCTCGACTTCAATTCACTCCGATTCCTCTCCCTCC 294
QY 181 TTGCGCACGCGCACACACGCGCACTCTGCTCTCTGCGCACCATTAATGATGCTGCC 240
DB 295 TTGCGCACGCGCACACACGCGCACTCTGCTCTCTGCGCACCATTAATGATGCTGCC 354
QY 241 GAACGCGACGCTCCCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 300
DB 355 GAACGCGACGCTCCCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 414
QY 301 TTAACCGATGAATCAGAAAGCATTCGAAGCTAGGGTTTCGAAACCGCGCAATTGGT 360
DB 415 TTAACCGATGAATCAGAAAGCATTCGAAGCTAGGGTTTCGAAACCGCGCAATTGGT 474

QY 361 TTCAGCGACGACGCTTAAATCAGCCGAGACAGATTCTTCAAGCTGCTCGAAACTCTG 420
DB 475 TTCAGCGACGACGCTTAAATCAGCCGAGACAGATTCTTCAAGCTGCTCGAAACTCTG 534
QY 421 TCTAATCTCGGCTCTAGAAAGAGTACATGAAGGCTCTTCTGATGATGAAGAGCTACA 480
DB 535 TCTAATCTCGGCTCTAGAAAGAGTACATGAAGGCTCTTCTGATGATGAAGAGCTACA 594
QY 481 GTCATCATGATGTTCTTGGGATTAAGGTTCTCGGGGCTCTCTGTGATTCAGAAAGGT 540
DB 595 GTCATCATGATGTTCTTGGGATTAAGGTTCTCGGGGCTCTCTGTGATTCAGAAAGGT 654
QY 541 GGTGAGACTGAGATGATCTTCTGGGTTGATGAGGCTCTGCTTAAGAGAGGTTGCTTAAG 600
DB 655 GGTGAGACTGAGATGATCTTCTGGGTTGATGAGGCTCTGCTTAAGAGAGGTTGCTTAAG 714
QY 601 TCGTTTAAGCAAGATGCTGTTTAAGTATGCGCTTCGCTTCCATGCTCGAGGGAT 660
DB 715 TCGTTTAAGCAAGATGCTGTTTAAGTATGCGCTTCGCTTCCATGCTCGAGGGAT 774
QY 661 GCTATGGCATTTGATCCACCTGATTTTATTAAGTATGATGATGTTGTTGAGAAAGCTTTG 720
DB 775 GCTATGGCATTTGATCCACCTGATTTTATTAAGTATGATGATGTTGTTGAGAAAGCTTTG 834
QY 721 AAGCTTTTACAG 780
DB 835 AAGCTTTTACAG 894
QY 781 GAGACTTTGGAAGAGATCACTCCCGCTTATGCTTGAAGTACTTGGCTTACCGCTTGT 840
DB 895 GAGACTTTGGAAGAGATCACTCCCGCTTATGCTTGAAGTACTTGGCTTACCGCTTGT 954
QY 841 GATGATTAACCTCGAGAAAGACTAAATGTTTAAAGCGTGTGCGGAATATTTTGTGCT 900
DB 955 GATGATTAACCTCGAGAAAGACTAAATGTTTAAAGCGTGTGCGGAATATTTTGTGCT 1014
QY 901 GTTGAAGAGGTGAGAGATCAGCTCTTGTGGGAGTTGAACCGGTGAAGAGTTATGAT 960
DB 1015 GTTGAAGAGGTGAGAGATCAGCTCTTGTGGGAGTTGAACCGGTGAAGAGTTATGAT 1074
QY 961 GAGGCGTTTTTAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
DB 1075 GAGGCGTTTTTAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1134
QY 1021 AATATTCACAGAGATCAATTTGAAGTTTGAAGTTGCACTTCTCTTGTGCTCAAGCT 1080
DB 1135 AATATTCACAGAGATCAATTTGAAGTTTGAAGTTGCACTTCTCTTGTGCTCAAGCT 1194
QY 1081 TTTATTTGTAAGAAAGCCACCTTTTACAGAGATGCTGATGAAGCAATTCAGCACTT 1140
DB 1195 TTTATTTGTAAGAAAGCCACCTTTTACAGAGATGCTGATGAAGCAATTCAGCACTT 1254
QY 1141 CAGGCTTAAGGTAATGCTATGAGATTCCTCGCATGTTGTAATGATACACGGAATTAAT 1200
DB 1255 CAGGCTTAAGGTAATGCTATGAGATTCCTCGCATGTTGTAATGATACACGGAATTAAT 1314
QY 1201 GAGATAGACTTCCGCTCTAGAAAGGAGACTGTGCACTGCTTATGAGCAAAAGTTGATGA 1260
DB 1315 GAGATAGACTTCCGCTCTAGAAAGGAGACTGTGCACTGCTTATGAGCAAAAGTTGATGA 1374
QY 1261 TGCCGTAATGTTGTTGGCTTTAGACAGTGAAGATTCACAAATATGAAATCCAGTATTT 1320
DB 1375 TGCCGTAATGTTGTTGGCTTTAGACAGTGAAGATTCACAAATATGAAATCCAGTATTT 1434
QY 1321 GAGTTTGTTTTGAAGAAATCAATCGTGAAGAAATGATATCCCTGGAATTAATGAA 1380
DB 1435 GAGTTTGTTTTGAAGAAATCAATCGTGAAGAAATGATATCCCTGGAATTAATGAA 1494
QY 1381 TTGTTGGAACCTGTTTGGAGAGGTTGCTTCTTAAGTTCAAGACACCAAAAGATTA 1440
DB 1495 TTGTTGGAACCTGTTTGGAGAGGTTGCTTCTTAAGTTCAAGACACCAAAAGATTA 1554

```

QY 1441 AAATTTAACTGGGGGACTACTATGATGATCTATGCTTTGATCTTGTGAAAGAGTG 1500
DB 1555 AAATTTAACTGGGGGACTACTATGATGATCTATGCTTTGATCTTGTGAAAGAGTG 1614
QY 1501 GAGGTAGTTCAGGGTCTCCTTTAGCTGCTGCACTATGCGCAAGATGAGGCGGAG 1560
DB 1615 GAGGTAGTTCAGGGTCTCCTTTAGCTGCTGCACTATGCGCAAGATGAGGCGGAG 1674
QY 1561 CATGTGAAAAGTACTAGTCTATGAGGCACTGCAAGAAAGTTTCTTCCGCTATACAGAT 1620
DB 1675 CATGTGAAAAGTACTAGTCTATGAGGCACTGCAAGAAAGTTTCTTCCGCTATACAGAT 1734
QY 1621 AGAAACTCGGCTGAACCCAGAGATGTGCAAGAGACGTGTATGATGATCTGTTGGT 1680
DB 1735 AGAAACTCGGCTGAACCCAGAGATGTGCAAGAGACGTGTATGATGATCTGTTGGT 1794
QY 1681 AACATGTAGGCGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
DB 1795 AACATGTAGGCGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1854
QY 1741 GAAAACTTTGAAACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800
DB 1855 GAAAACTTTGAAACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1914
QY 1801 GAACTACTGTTGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1860
DB 1915 GAACTACTGTTGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1974
QY 1861 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1920
DB 1975 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2034
QY 1921 AGCTCATCTTTTCAACGCAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1980
DB 2035 AGCTCATCTTTTCAACGCAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2094
QY 1981 GGGTCAGTCAGAGCTGACGATTCAGAGACCTTCCAGAGATGATGATGATGATGATGATGAT 2040
DB 2095 GGGTCAGTCAGAGCTGACGATTCAGAGACCTTCCAGAGATGATGATGATGATGATGATGAT 2154
QY 2041 AATATAGTATCCAGTGGCAGAAATTAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
DB 2155 AATATAGTATCCAGTGGCAGAAATTAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2214
QY 2101 GAAATGTTTACAGAGGTTTGGATGGGCGAATGCTGAAGATTGGAATGGAATGGAATGGAAT 2160
DB 2215 GAAATGTTTACAGAGGTTTGGATGGGCGAATGCTGAAGATTGGAATGGAATGGAATGGAAT 2274
QY 2161 GAAACTGCGCAGCTTGGGTTGATGATTAATTAACATGTTGAAATCTATCTGTTGACAGT 2220
DB 2275 GAAACTGCGCAGCTTGGGTTGATGATTAATTAACATGTTGAAATCTATCTGTTGACAGT 2334
QY 2221 GTGACAGTCTCAGCGAGATGAAACCCGTGCTCTGCTGGAAGCAACTCTGGAAGAGTCTGCT 2280
DB 2335 GTGACAGTCTCAGCGAGATGAAACCCGTGCTCTGCTGGAAGCAACTCTGGAAGAGTCTGCT 2394
QY 2281 TGTCTATCTGATTTGGTTCATCTCAGAAAACAATGCTATGATGATGATGATGATGATGATGAT 2340
DB 2395 TGTCTATCTGATTTGGTTCATCTCAGAAAACAATGCTATGATGATGATGATGATGATGATGAT 2454
QY 2341 AGATACGAAGTTTCTGCTCAGAGTCAAGGTTGAAATCACTGAAGGCTCTGTTCTTCA 2400
DB 2455 AGATACGAAGTTTCTGCTCAGAGTCAAGGTTGAAATCACTGAAGGCTCTGTTCTTCA 2514
QY 2401 TCATTA 2406
DB 2515 TCATTA 2520

```

RESULT 6
ADJ38130
ID ADJ38130 standard; DNA; 3667 BP.
XX

```

AC ADJ38130;
XX
XX 06-MAY-2004 (first entry)
DB Arabidopsis thaliana Arc6-1 genomic DNA SegID2.
XX
XX prokaryotic type; plastid division; Ftn2, ARC6, ARC5, Fzo; plant cell;
XX agronomic; horticultural; crop plant; ornamental plant; woody plant;
XX herbicide target; ds.
XX
XX Arabidopsis thaliana.
XX
XX WO2004001003-A2.
XX
XX 31-DEC-2003.
XX
XX 20-JUN-2003; 2003WO-US019536.
XX
XX 20-JUN-2002; 2002US-0390140P.
XX
XX 09-AUG-2002; 2002US-0402242P.
XX
XX 20-JUN-2003; 2003US-00600070.
XX
XX (UNMS ) UNIV MICHIGAN STATE.
XX
XX Oseleyoung KW, Vilha S, Koksharova OA, Gao H;
XX
XX WPI: 2004-082486/08.
XX
XX P-PSDB; ADJ38202.
XX
XX New isolated Ftn2, ARC5 and/or Fzo-like nucleic acid sequences, useful
XX for further characterizing plastid division in plant cells, and in
XX PT varying agronomic and horticultural characteristics of economically
XX PT important plants.
XX
XX Claim 1; SEQ ID NO 2; 287bp; English.
XX
XX
XX This invention relates to novel prokaryotic type or plastid division and
XX CC related genes and proteins. In particular, the invention relates to novel
XX CC Ftn2 (ARC6), ARC5 and Fzo-like genes and polypeptides. The methods and
XX CC compositions of the present invention are useful for further
XX CC characterizing plastid division in plant cells, in order to vary
XX CC agronomic and horticultural characteristics of economically important
XX CC plants, such as crop, ornamental and woody plants. They can also be used
XX CC as herbicide targets. The present sequence is that of a DNA sequence
XX CC which is related to the invention.
XX
XX
XX Sequence 3667 BP, 983 A; 670 C; 842 G; 1172 T; 0 U; 0 Other;
XX
XX
XX Query Match 71.7%; Score 1724; DB 12; Length 3667;
XX Best Local Similarity 79.2%; Pred. No. 0;
XX Matches 2406; Conservative 0; Mismatches 0; Indels 632; Gaps 5;
QY 1 ATGAAAGCTTGAAGTCACTGCGGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60
DB 481 ATGAAAGCTTGAAGTCACTGCGGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
QY 61 CCGGAGACGAAAGCTCCGAGCGTACGCAACCAACCTTAAACATATGCTGCTGCTGCTGCTGCTGCTG 120
DB 541 CCGGAGACGAAAGCTCCGAGCGTACGCAACCAACCTTAAACATATGCTGCTGCTGCTGCTGCTGCTG 600
QY 121 AAATGAGCGCAGCGTCTCTCTCCGACTTCAATTTCACTCCGATTCCTCCTCCTCCTCCTCCTCCTC 180
DB 601 AAATGAGCGCAGCGTCTCTCTCCGACTTCAATTTCACTCCGATTCCTCCTCCTCCTCCTCCTCCTC 660
QY 181 TTGCGCACCGGCACACCAACCGCCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
DB 661 TTGCGCACCGGCACACCAACCGCCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
QY 241 GAACGCCAGTCCCGCATCCCATGATTTCTTACACAGGATTAAGAGCTCAAAACATTTTC 300
DB 721 GAACGCCAGTCCCGCATCCCATGATTTCTTACACAGGATTAAGAGCTCAAAACATTTTC 780
QY 301 TTAACCGATGGAATCAGAAAGACATTCGAAAGCTAGGGTTTGGAAACGCGCAATTCCGT 360

```

```
|||||
Db 781 TTAACCATGGAAATCAGAAAGACATTCGAGCTGAGGTTTGAAACCGCCGCAATTCGGT 840
Qy 361 TTAGGACACGCGTTTAATCAGCCGAGACAGATTCTTCAAGCTGCTGGAAACTG 420
Db 841 TTACGACACGCGTTTAATCAGCCGAGACAGATTCTTCAAGCTGCTGGAAACTG 900
Qy 421 TCTAATCCTCGGCTAGAGAGAGATAGAAAGTCTTCTGATGATGAAGAGCTACA 480
Db 901 TCTAATCCTCGGCTAGAGAGAGATAGAAAGTCTTCTGATGATGAAGAGCTACA 960
Qy 481 GTATCATCTGATGTTCTTGGGAT----- 504
Db 961 GTCATCATGATGTTCTTGGGATGATGATATTCGATTCGGAATATAAAGTTTCTC 1020
Qy 505 -----AAGTTCTGGGGC 518
Db 1021 GTTTTAAATTTTCATTAATGGATTAAGAGAACTTTTATCTAGTGAAGGTTCTGGGGC 1080
Qy 519 TCTCTGTGATATGACAGAAAGGTGGTGAGACTGAGATAGTTCTTCGGGTTGGTGAAGCTCT 578
Db 1081 TCTCTGTGATATGACAGAAAGGTGGTGAGACTGAGATAGTTCTTCGGGTTGGTGAAGCTCT 1140
Qy 579 GCTTAAGAGAGGTTGCTTAAGTCTTTAAGCAAGATGTGTTTAACTTATGCGCTTGC 638
Db 1141 GCTTAAGAGAGGTTGCTTAAGTCTTTAAGCAAGATGTGTTTAACTTATGCGCTTGC 1200
Qy 639 GTTTCGTGATGTCGAGGAGATGCTATGAGCATGAGATCAACGTATTTTATTACTGCTTA 698
Db 1201 GTTTCGTGATGTCGAGGAGATGCTATGAGCATGAGATCAACGTATTTTATTACTGCTTA 1260
Qy 699 TGAATTTGTAGAGAGCTTTGAAGCTTTTAC----- 730
Db 1261 TGAATTTGTAGAGAGCTTTGAAGCTTTTACAGGTATGTTGACTGCTTTGTAATTTG 1320
Qy 731 -----A 731
Db 1321 ACGAGCGTTGGCTTTATAGAACTTCTGATTTGATTACTTTGATTTGAGTCTTGTGTA 1380
Qy 732 GGAGAGAGAGAGAGTACCTTGCACCGGATTTACGTGCACAAATTTGAGACTTTGGA 791
Db 1381 GGAGAGAGAGAGAGTACCTTGCACCGGATTTACGTGCACAAATTTGAGACTTTGGA 1440
Qy 792 AGAGATCACTCCGCTTATGTCCTTGAGACTTATGAGCTTACCGCTGGTGAATTAACG 851
Db 1441 AGAGATCACTCCGCTTATGTCCTTGAGACTTATGAGCTTACCGCTGGTGAATTAACG 1500
Qy 852 TGGCAAAAGACTTAATGCTTTAAGCGGTGCGGAATTTTGTGTCTGTGAGAGAG 911
Db 1501 TGGCAAAAGACTTAATGCTTTAAGCGGTGCGGAATTTTGTGTCTGTGAGAGAG 1560
Qy 912 TGGAGATCAAGCTTTGTTGGGGGTTTGAACCGTGAAGAGTTTATGATGAGGCGTTT 971
Db 1561 TGGAGATCAAGCTTTGTTGGGGGTTTGAACCGTGAAGAGTTTATGATGAGGCGTTT 1620
Qy 972 ACGAGATCAAGCTGCTGAGC----- 991
Db 1621 ACGAGATCAAGCTGCTGAGCTATACAGTTAGATACCTTTTAAATTTCTTTAGCA 1680
Qy 992 -----AGTTGATCTTTTGTGA 1008
Db 1681 TGATATACTTTAGGTTCTCATTTTAATGATGTTGTGTGTGTTGTTGTTGTTGTA 1740
Qy 1009 GCTACCCCAAGCAATATTTCCAGAGAGTCAATTTGAAGTTTACGAAGTTGACTGCTT 1068
Db 1741 GCTACCCCAAGCAATATTTCCAGAGAGTCAATTTGAAGTTTACGAAGTTGACTGCTT 1800
Qy 1069 GTGCTCAAGCTTTTATTTGTTAGAGAGCAACCTTTTACAGATGCTGATTAACAATTC 1128
Db 1801 GTGCTCAAGCTTTTATTTGTTAGAGAGCAACCTTTTACAGATGCTGATTAACAATTC 1180
Qy 1129 CAGCACTTACAGAGGCTAAGTAAATGCTATGAGATTCCTGCGAGTGTGTATGATACA 1188
|||||
Db 1861 CAGCACTTACAGAGGCTAAGTAAATGCTATGAGATTCCTGCGAGTGTGTATGATACA 1920
Qy 1189 CGGAATATAATTTGGAGATATAGACTTCGGCTCTAAGAAAGGGAATCTGTGCACTTAAAGG 1248
Db 1921 CGGAATATAATTTGGAGATATAGACTTCGGCTCTAAGAAAGGGAATCTGTGCACTTAAAGG 1980
Qy 1249 AAAGTTGATGAATGCGGATATGCTTGGGCTTGAACAGTGAAGATTCACAAATATAGGAAT 1308
Db 1981 AAAGTTGATGAATGCGGATATGCTTGGGCTTGAACAGTGAAGATTCACAAATATAGGAAT 2040
Qy 1309 CCAGCTATTTGAGATTTGTTTGGAGAAATCAATCTGATGACATGATGATCTCCCT 1368
Db 2041 CCAGCTATTTGAGATTTGTTTGGAGAAATCAATCTGATGACATGATGATCTCCCT 2100
Qy 1369 GGAATATGCAAAATTTGTTGGAACCTGGTTGGCAGGGGTTGCTTCTCTAGTTTCAGAGAC 1428
Db 2101 GGAATATGCAAAATTTGTTGGAACCTGGTTGGCAGGGGTTGCTTCTCTAGTTTCAGAGAC 2160
Qy 1429 ACCAAAGATTAATAATTTAAATCTCGGGACTACTATGATGATCTTATGTTTGAATTAAC 1488
Db 2161 ACCAAAGATTAATAATTTAAATCTCGGGACTACTATGATGATCTTATGTTTGAATTAAC 2220
Qy 1489 TTGGAAGAGTGAAGGATGTTGAGGGTCTCTTTAGCTGCTGCTGCAACTATGCGAAG 1548
Db 2221 TTGGAAGAGTGAAGGATGTTGAGGGTCTCTTTAGCTGCTGCTGCAACTATGCGAAG 2280
Qy 1549 ATTGAGCCGAGCATGTAAGCTAGTCTATGCAAGCACTGAGAAAGTTTCTCTCC 1608
Db 2281 ATTGAGCCGAGCATGTAAGCTAGTCTATGCAAGCACTGAGAAAGTTTCTCTCC 2340
Qy 1609 CGCTATACAGATAGAAATCTCGGCTGAACCCAGAGATGTGCAAGACAGTTTTAACTGTA 1668
Db 2341 CGCTATACAGATAGAAATCTCGGCTGAACCCAGAGATGTGCAAGACAGTTTTAACTGTA 2400
Qy 1669 GATTCCTGTTGTAACAATGAGGCGGATGAGGATGAGGCTGCTTATTTGAGAGAGCT 1728
Db 2401 GATTCCTGTTGTAACAATGAGGCGGATGAGGATGAGGCTGCTTATTTGAGAGAGCT 2460
Qy 1729 GTAAGACCCCTGTGAAACCTTTGAAACTAATGATTAATGCAATTTGAGCTGGGCTCTCAGAG 1788
Db 2461 GTAAGACCCCTGTGAAACCTTTGAAACTAATGATTAATGCAATTTGAGCTGGGCTCTCAGAG 2520
Qy 1789 AGTAGCGTTGATGAACCTACTGTTGAATGTCCGTTGCTGATATGTTAAAGAGGCAAGT 1848
Db 2521 AGTAGCGTTGATGAACCTACTGTTGAATGTCCGTTGCTGATATGTTAAAGAGGCAAGT 2580
Qy 1849 GTGAAGATCCTACTGCTGCTGTGCGAATTTGACTATTTCACTGTTCAAGCGAAGATAT 1908
Db 2581 GTGAAGATCCTACTGCTGCTGTGCGAATTTGACTATTTCACTGTTCAAGCGAAGATAT 2640
Qy 1909 TTTCTTAAAGACAGCTCATCTTTTCAACGCAAGATATGTTTCTTCTATGGAATCTGAT 1968
Db 2641 TTTCTTAAAGACAGCTCATCTTTTCAACGCAAGATATGTTTCTTCTATGGAATCTGAT 2700
Qy 1969 GTGCGTACCA----- 1978
Db 2701 GTGCGTACCAATAGATTAATGATGAATTTTCAATATATCTGATTCCTCAAAATA 2760
Qy 1979 ----- 1978
Db 2761 TGCTTTTGTGAGCTTAAGACATATGTTCCACTTAATACATGTCCTGATTTTGTATACC 2820
Qy 1979 ----- 1978
Db 2821 AAGATTACAAGTTGCTGATTAATTTCACTAATTAATGCTGCTGAATTTTGTATCAAA 2880
Qy 1979 ----- 1978
Db 2881 CTGTAGACAGAAATGTAATTTCACTCTCAACATTTCTGTTAGAAATACGATAGATTAAG 2940
Qy 1979 -----T 1979
Db 2941 AAGATTGCTTAAATGTTGCTTGTGCAACTTTTCTTTCCTGATTTTCTTTTGCATTTT 3000
```

QY 1980 AGGGTCAGTCAGAGCTGACGATTGAGAGCACTTCCAGATGAGTGTAGAGCTGCA 2039
DB 3001 AGGGTCAGTCAGAGCTGACGATTGAGAGCACTTCCAGATGAGTGTAGAGCTGCA 3060
QY 2040 GAATATGATGATCCAGATGAGAGGATTAAGTCTGCGCTTTGGCCCTGATCAGCGCAT 2099
DB 3061 GAATATGATGATCCAGATGAGAGGATTAAGTCTGCGCTTTGGCCCTGATCAGCGCAT 3120
QY 2100 AGAATGTTACCGAG-----AGATTTGATGGC 2113
DB 3121 AGAATGTTACCGAGAGGATTAATCTAATCAATGATGTTGATAAACTGT 3180
QY 2114 -----AGATTTGATGGC 2128
DB 3181 TGGACATGATTAATGCTGCTGCTTTGATTTCTGTTATTTATGATTTGATGGC 3240
QY 2129 GAATGCTGAAGATTTGAGTCTGACAGAGCTGAACTGCGAGCTTGGGTTGTTATG 2188
DB 3241 GAATGCTGAAGATTTGAGTCTGACAGAGCTGAACTGCGAGCTTGGGTTGTTATG 3300
QY 2189 ATTATGACTGTTGAATCTATCTGTGACAGTGTGACAGTCTGAGATGGAACCCGTG 2248
DB 3301 ATTATGACTGTTGAATCTATCTGTGACAGTGTGACAGTCTGAGATGGAACCCGTG 3360
QY 2249 CTCTGCTGAGAGCACTCTGAGAGTCTGTGCTATCTGATTTGTTCAATCCAGAA 2308
DB 3361 CTCTGCTGAGAGCACTCTGAGAGTCTGTGCTATCTGATTTGTTCAATCCAGAA 3420
QY 2309 ACAATGCTACTGATGTGACAGACTTACACAAAGATAGCAAGTTTCTGCTCAAGTCAG 2368
DB 3421 ACAATGCTACTGATGTGACAGACTTACACAAAGATAGCAAGTTTCTGCTCAAGTCAG 3480
QY 2369 GGTGGAATCACTGAGAGCTCTGTTCTGCAATCAATA 2406
DB 3481 GGTGGAATCACTGAGAGCTCTGTTCTGCAATCAATA 3518

RESULT 7
ADJ38136
ID ADJ38136 standard; DNA; 3667 BP.

XX ADJ38136;
AC
XX
DT 06-MAY-2004 (first entry)
XX
DE Arabidopsis thaliana AtFtn2 genomic DNA SegID10.
XX
KW prokaryotic type; plastid division; Ftn2; ARC6; ARC5; Fzo; plant cell;
KW agronomic; horticultural; crop plant; ornamental plant; woody plant;
KW herbicide target; ds.
XX
OS Arabidopsis thaliana.
XX
PN WO2004001003-A2.
XX
PD 31-DEC-2003.
XX
PF 20-JUN-2003; 2003WO-US019536.
XX
PR 20-JUN-2002; 2002US-0390140P.
XX
PR 09-AUG-2002; 2002US-0402242P.
XX
PR 20-JUN-2003; 2003US-00600070.
XX
PA (UNMS) UNIV MICHIGAN STATE.
XX
PI Oosteryoung KW, Vitha S, Koksharova OA, Gao H,
DR WPI; 2004-082486/08.
XX P-PSDB; ADJ38203.
XX
PT New isolated Ftn2, ARCS and/or Fzo-like nucleic acid sequences, useful
PT for further characterizing plastid division in plant cells, and in

PT varying agronomic and horticultural characteristics of economically
PT important plants.
XX
XX
PS Claim 1; SEQ ID NO 10; 287bp; English.

CC This invention relates to novel prokaryotic type or plastid division and
CC related genes and proteins. In particular, the invention relates to novel
CC Ftn2 (ARC6), ARCS and Fzo-like genes and polypeptides. The methods and
CC compositions of the present invention are useful for further
CC characterizing plastid division in plant cells, in order to vary
CC agronomic and horticultural characteristics of economically important
CC plants, such as crop, ornamental and woody plants. They can also be used
CC as herbicide targets. The present sequence is that of a DNA sequence
CC which is related to the invention.
XX
XX

SEQ Sequence 3667 BP; 982 A; 669 C; 843 G; 1173 T; 0 U; 0 Other;

Query Match 71.5%; Score 1720.8; DB 12; Length 3667;
Best Local Similarity 79.1%; Pred. No. 0;
Matches 2404; Conservative 0; Mismatches 2; Indels 632; Gaps 5;

QY 1 ATGGAAGCTCTGAGTCAGTGGGCAATGGTCTCCGCCATTCATTTAGCGATTAACA 60
DB 491 ATGGAAGCTCTGAGTCAGTGGGCAATGGTCTCCGCCATTCATTTAGCGATTAACA 540
QY 61 CCGGCGAGAGCAAGCTCCGAGTAGCACAACCTTACAACTATCTGCTCCGACAG 120
DB 541 CCGGCGAGAGCAAGCTCCGAGTAGCACAACCTTACAACTATCTGCTCCGACAG 600
QY 121 AAATGGCGCAACCGTCTTCTCCGACTTCAATTTACCTCCGATTCCTCTCTCC 180
DB 601 AAATGGCGCAACCGTCTTCTCCGACTTCAATTTACCTCCGATTCCTCTCTCC 660
QY 181 TTGCGACACCGCACACACACCGCACTGCTCTCTGCGACCATATGATGATGCTCC 240
DB 661 TTGCGACACCGCACACACACCGCACTGCTCTCTGCGACCATATGATGATGCTCC 720
QY 241 GAAAGCCAGTCCCATCCCATTTGATTTCTACAGAGTATTAGAGCTCAACACATTTG 300
DB 721 GAAAGCCAGTCCCATCCCATTTGATTTCTACAGAGTATTAGAGCTCAACACATTTG 780
QY 301 TTAACCGATGGAATCAGAAAGCATTTGAGCTGAGGTTTGAACCGCGCAATTCGGT 360
DB 781 TTAACCGATGGAATCAGAAAGCATTTGAGCTGAGGTTTGAACCGCGCAATTCGGT 840
QY 361 TTGACGACGACGCTTTAATGACCGGAGACAGATTTCTCAAGCTGCTTGGAACTGTG 420
DB 841 TTGACGACGACGCTTTAATGACCGGAGACAGATTTCTCAAGCTGCTTGGAACTGTG 900
QY 421 TCTAATCCTCGGTCTAGAGAGATGATCAATGAAGTCTTCTTGATGATGAAGAGTACA 480
DB 901 TCTAATCCTCGGTCTAGAGAGATGATCAATGAAGTCTTCTTGATGATGAAGAGTACA 960
QY 481 GTCATCATGATGTTCTTGGGAT----- 504
DB 961 GTCATCATGATGTTCTTGGGATGAGTATTCGATTTGGAATATAAGTTTCTTC 1020
QY 505 -----AAGTTTCCGCGGC 518
DB 1021 GTTTAAATTCATGAATGATTAAGAGAGAACTTTATCTAGTGAAGGTTCTCGGGC 1080
QY 519 TCTCTGATGATTTGAGAGAGTGTGAGACTGATAGTATCTTCCGGTGTGAGGCTCT 578
DB 1081 TCTCTGATGATTTGAGAGAGTGTGAGACTGATAGTATCTTCCGGTGTGAGGCTCT 1140
QY 579 GCTTAAGAGAGGTTGCTTAAGTCTTTAAGCAAGATGATGATTTAGTATGCGCTTCG 638
DB 1141 GCTTAAGAGAGGTTGCTTAAGTCTTTAAGCAAGATGATGATTTAGTATGCGCTTCG 1200
QY 639 GTTCTGATGATGCTCGAGGAGTGTATGATGATGATGATGATGATGATGATGATGAT 698
DB 1201 GTTCTGATGATGCTCGAGGAGTGTATGATGATGATGATGATGATGATGATGATGAT 1260

QY 699 TGAGTTGTTGAGAGCTTGAACCTTTAC----- 730
Db 1261 TGAGTTGTTGAGAGCTTTGAACCTTTTACAGTAAGTTGACTTGTGTAATTG 1320
QY 731 -----A 731
Db 1321 ACGACGTTGGCTTTATAGAACTTTCTTGATTGTACTTGTATTGAGCTTGTGTA 1380
QY 732 GGAGGAGAGAGCAAGTAAGCTTGCACCGGATTTAGTGCACAAAATTGATGAGACTTTGA 791
Db 1381 GGAGGAGAGAGCAAGTAAGCTTGCACCGGATTTAGTGCACAAAATTGATGAGACTTTGA 1440
QY 792 AGAGATCACTCCGCGTTATGCTTGTGAGAGCTACTTGGCTTACCGCTTGTGATGATTAAGC 851
Db 1441 AGAGATCACTCCGCGTTATGCTTGTGAGAGCTACTTGGCTTACCGCTTGTGATGATTAAGC 1500
QY 852 TCGCAAAAGACTTAATGCTTTAAGCGGTGTGCGGAATATTTTGTGTCTGTGAGAGAG 911
Db 1501 TCGCAAAAGACTTAATGCTTTAAGCGGTGTGCGGAATATTTTGTGTCTGTGAGAGAG 1560
QY 912 TGGAGCATCAGCTCTTGTGGGGGTTTGACCCGTGAGAGTTTATGAAAGGCGTTT 971
Db 1561 TGGAGCATCAGCTCTTGTGGGGGTTTGACCCGTGAGAGTTTATGAAAGGCGTTT 1620
QY 972 ACGAATGACAGCTGCTGAGC----- 991
Db 1621 ATGATGACAGCTGCTGAGCAGATPAACGTTAGATCCTTTTAAATTTCTTTAGCA 1680
QY 992 -----AGTTGATCTTTTGTGA 1008
Db 1681 TGATATAACTTTAGGTTCTCATTTTAAATGTAATGTTGTGTGATGATGATCTTTTGTGA 1740
QY 1009 GCTACCCCAAGCAATATTCACAGAGTCAATTGAAAGTTTACGAAGTTGCACCTTCT 1068
Db 1741 GCTACCCCAAGCAATATTCACAGAGTCAATTGAAAGTTTACGAAGTTGCACCTTCT 1800
QY 1069 GTGGCTCAAGCTTTTATGTTAAGAACCACTTTTACAGAGTCTGATPAAGCAATTC 1128
Db 1801 GTGGCTCAAGCTTTTATGTTAAGAACCACTTTTACAGAGTCTGATPAAGCAATTC 1860
QY 1129 CAGCAACTTCACAGAGCTTAAGTAAATGCTATGAGATTCCTGCGATGTTGTAATAC 1188
Db 1861 CAGCAACTTCACAGAGCTTAAGTAAATGCTATGAGATTCCTGCGATGTTGTAATAC 1920
QY 1189 CGGAATTAATTTGGAGATAGACTTCGCTCTAGAAAGGGGACTCTGCACTGCTTAAAGC 1248
Db 1921 CGGAATTAATTTGGAGATAGACTTCGCTCTAGAAAGGGGACTCTGCACTGCTTAAAGC 1980
QY 1249 AAAAGTTGATGAATGCCGTATGTGTGGCTTGAACAGTGAGAGATTCAAAATPAAGAA 1308
Db 1981 AAAAGTTGATGAATGCCGTATGTGTGGCTTGAACAGTGAGAGATTCAAAATPAAGAA 2040
QY 1309 CCAGCTATTTGAGAGTTTGTGTTGAGAAATTCAAATCGTGAATGACAAATGATCTCCCT 1368
Db 2041 CCAGCTATTTGAGAGTTTGTGTTGAGAAATTCAAATCGTGAATGACAAATGATCTCCCT 2100
QY 1369 GGACTATGCAAAATTTGTTGAAAACCTGTGGCAGGGGTTGTCTTCTAGGTTCCAGAGC 1428
Db 2101 GGACTATGCAAAATTTGTTGAAAACCTGTGGCAGGGGTTGTCTTCTAGGTTCCAGAGC 2160
QY 1429 ACCAAAGATAAAAATTTAAATCTCGGGGACTACTATGATGATCTATGAGTTTGAATTAC 1488
Db 2161 ACCAAAGATAAAAATTTAAATCTCGGGGACTACTATGATGATCTATGAGTTTGAATTAC 2220
QY 1489 TTGGAAGAGTGGAGTACTTCAAGGTTTCTCTTTAAGCTGCTGCAACTATGCGAAGG 1548
Db 2221 TTGGAAGAGTGGAGTACTTCAAGGTTTCTCTTTAAGCTGCTGCAACTATGCGAAGG 2280
QY 1549 ATTGAGCGGAGCATGTGAAAAGCTAGTGTATGAGGAGCCTGCAAAAAGTTTCTTCC 1608
Db 2281 ATTGAGCGGAGCATGTGAAAAGCTAGTGTATGAGGAGCCTGCAAAAAGTTTCTTCC 2340
QY 1609 CGCTATACAGATAGAAATCTGGCTGAACCAAGATGTGCAAGAGACAGTGTTTAGTGA 1668

Db 2341 CGCTATACAGATAGAAATCTGGCTGAACCAAGATGTGCAAGAGACAGTGTTTAGTGA 2400
QY 1669 GATCTGTGGTAAACAATGTAAGCCGTGATGAGCTGTGTCTTTATTTGAGAGACT 1728
Db 2401 GATCTGTGGTAAACAATGTAAGCCGTGATGAGCTGTGTCTTTATTTGAGAGACT 2460
QY 1729 GTAAGACCTCTGAAAACCTTTGAAAACATAATGATTAATTCAGAGCTGGGGTCTCAGAG 1788
Db 2461 GTAAGACCTCTGAAAACCTTTGAAAACATAATGATTAATTCAGAGCTGGGGTCTCAGAG 2520
QY 1789 AGTAGCGTTGATGAACCTACTGTTGAAAATGTCGTTGCTGATATGTTAAAGAGGCAACT 1848
Db 2521 AGTAGCGTTGATGAACCTACTGTTGAAAATGTCGTTGCTGATATGTTAAAGAGGCAACT 2580
QY 1849 GTGAAGATCCAGTGTGCGGTGGTGGCAATGGAATGACATGATTCACGTTCAGCAGGAATAT 1908
Db 2581 GTGAAGATCCAGTGTGCGGTGGTGGCAATGGAATGACATGATTCACGTTCAGCAGGAATAT 2640
QY 1909 TTTCTTAAAGCAGCTCATCTTTTCAACGCAAGATATGTTTCTTCTATGGAATCTGAT 1968
Db 2641 TTTCTTAAAGCAGCTCATCTTTTCAACGCAAGATATGTTTCTTCTATGGAATCTGAT 2700
QY 1969 GTGCTTACCA----- 1978
Db 2701 GTGCTTACCAATAGGATATTAATGATGCAATTTTCAATATCTGATTCCTCAAAATA 2760
QY 1979 ----- 1978
Db 2761 TGCTTGTGTTGTGACTTAAGAACATAGTCTCCACTTAATACATGTCCCAAAAGTTGTAAC 2820
QY 1979 ----- 1978
Db 2821 AAGATTAACAAGTTGCTGAGTAATTTCACTAATTAATGCTGCTGAATTTTTGATCAAA 2880
QY 1979 ----- 1978
Db 2881 CTGTAGACAGAAATGTAAATTTCACTCTCAACTTCTGTTTAAATPAAGTAAGATTAG 2940
QY 1979 ----- 1979
Db 2941 AGATTCCTTAATGTTGCTGCTTGTTCCAAATTTTCTTCTGATTTTTTCTTTTCAATTT 3000
QY 1980 AGGCTCAGTCAGACCTCAGCAATTCAGAAAGCACTTCCAGAAATGAGATCTAGACCTGAGA 2039
Db 3001 AGGCTCAGTCAGACCTCAGCAATTCAGAAAGCACTTCCAGAAATGAGATCTAGACCTGAGA 3060
QY 2040 GAATATGATATCCAAAGTGGCAGAAAGATTAAATCTCTGGCTTTTGGGCTGATCACCGCAT 2099
Db 3061 GAATATGATATCCAAAGTGGCAGAAAGATTAAATCTCTGGCTTTTGGGCTGATCACCGCAT 3120
QY 2100 AGAATGTATACAG----- 2113
Db 3121 AGAATGTATACAGAGGTAGGGAATTAATCTACAAATTCAAATGATGTGAAAACTGT 3180
QY 2114 -----AGTTTGTGATGGC 2128
Db 3181 TGAGCATGATTAATGATGCTGCTGCTGTTGATTTCTGTTAATTAAGTTTGGATGGGC 3240
QY 2129 GAATGCTGAAAGATTGGACTGACAGAGAGCTGAAATCTGGCAGCTTGGGTTGTTATG 2188
Db 3241 GAATGCTGAAAGATTGGACTGACAGAGAGCTGAAATCTGGCAGCTTGGGTTGTTATG 3300
QY 2189 ATTATTAAGCTGTAACCTATCTGTTTGAACAGTGAACAGTCTGAGCATGGAAGCCCGTG 2248
Db 3301 ATTATTAAGCTGTAACCTATCTGTTTGAACAGTGAACAGTCTGAGCATGGAAGCCCGTG 3360
QY 2249 CTCTGTGGAAGCAACTCTGGAGAGTCTGCTTCTATCTGATTTGATTCATCAAGAA 2308
Db 3361 CTCTGTGGAAGCAACTCTGGAGAGTCTGCTTCTATCTGATTTGATTCATCAAGAA 3420
QY 2309 ACAATGCTACTGATGCTGAACCTTACACAAAGATACGAAGTTTCTGGTCCAAGTCAAG 2368

DB 3421 ACAATGCTACTGATGTCAGAACCTTACACAAAGATACGAAGTTTCTGGTCCAAAGTAG 3480
QY 2369 GGTGGAATCACTGAGGCTGTCTTTCATCATTA 2406
DB 3481 GGTGGAATCACTGAGGCTGTCTTTCATCATTA 3518
RESULT 8
ADJ38212/c
ID ADJ38212 standard; cDNA; 561 BP.
XX ADJ38212;
AC
XX
XX 06-MAY-2004 (first entry)
DE Plastid division-related Arc6 orthologue cDNA 5.
XX
XX prokaryotic type; plastid division; Ftn2; ARC6; ARC5; Fzo; plant cell;
KM agronomic; horticultural; crop plant; ornamental plant; woody plant;
KW herbicide target; gene; ss.
XX
XX Arabidopsis thaliana.
OS
XX
XX WO2004001003-A2.
PN
XX 31-DEC-2003.
XX
XX 20-JUN-2003; 2003WO-US019536.
PF
XX
XX 20-JUN-2002; 2002US-0390140P.
PR 09-AUG-2002; 2002US-0402242P.
PR 20-JUN-2003; 2003US-00600070.
XX
XX (UNMS) UNIV MICHIGAN STATE.
PA
XX
XX Osteryoung KW, Vitna S, Koksharova OA, Gao H;
PI WPI; 2004-082486/08.
XX
XX
XX New isolated Ftn2, ARC5 and/or Fzo-like nucleic acid sequences, useful
PT for further characterizing plastid division in plant cells, and in
PT varying agronomic and horticultural characteristics of economically
PI important plants.
XX
XX
XX Disclosure; Fig 8; 287pp; English.
PS
XX
XX This invention relates to novel prokaryotic type or plastid division and
CC related genes and proteins. In particular, the invention relates to novel
CC Ftn2 (ARC6), ARC5 and Fzo-like genes and polypeptides. The methods and
CC compositions of the present invention are useful for further
CC characterizing plastid division in plant cells, in order to vary
CC agronomic and horticultural characteristics of economically important
CC plants, such as crop, ornamental and woody plants. They can also be used
CC as herbicide targets. The present sequence is a cDNA sequence which is
CC related to the invention.
XX
XX
SQ Sequence 561 BP; 167 A; 136 C; 107 G; 148 T; 0 U; 3 Other;
Query Match 20 0%; Score 482; DB 12; Length 561;
Best Local Similarity 99.4%; Pred. No. 1.5e-130;
Matches 482; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1922 GGTCACTTTTCAAGCAAGATATGTTCTTCTATGAAATCTGATGCGTACCATAG 1981
DB 561 GGTCACTTTTCAAGCAAGATATGTTCTTCTATGAAATCTGATGCGTACCATAG 502
QY 1992 GGTCACTGAGCTGACGATTCAGAGCACTTCCAGAAATGATCTAGAGACTGACAGA 2041
DB 501 GGTCACTGAGCTGACGATTCAGAGCACTTCCAGAAATGATCTAGAGACTGACAGA 442
QY 2042 ATATGTTATCCAAAGGAGCAAGATTAAGTCTCTGCGTTTGGGCTGATCCCGCATAG 2101
DB 441 ATATGTTATCCAAAGGAGCAAGATTAAGTCTCTGCGTTTGGGCTGATCCCGCATAG 382

QY 2102 AAATGTTACGAGAGTTTGGATGCGCAATGCTGAAAGATTGGACTGACAGAGCTG 2161
DB 381 AAATGTTACGAGAGTTTGGATGCGCAATGCTGAAAGATTGGACTGACAGAGAGCTG 322
QY 2162 AAATGCGCAGCTTGGGTTGTTATGATTATATACCTGTGAAACTATCTGTTGACAGTG 2221
DB 321 AAATGCGCAGCTTGGGTTGTTATGATTATATACCTGTGAAACTATCTGTTGACAGTG 262
QY 2222 TGACAGTCTGACGAGATGAAACCGGTGCTGTGTTGGAAGCAACTGTGAGAGAGTGGCTT 2281
DB 261 TGACAGTCTGACGAGATGAAACCGGTGCTGTGTTGGAAGCAACTGTGAGAGAGTGGCTT 202
QY 2282 GGTATCTGATTTGGTTTCATCCAGAAACAAATGCTACTGATGTCAGAACTTACACAA 2341
DB 201 GGTATCTGATTTGGTTTCATCCAGAAACAAATGCTACTGATGTCAGAACTTACACAA 142
QY 2342 GATACGAAGTTTTCGTGTCGAAGTCAGGTTGGAANAATCACTGAAAGGCTGTCTTGCAAT 2401
DB 141 GATACGAAGTTTTCGTGTCGAAGTCAGGTTGGAANAATCACTGAAAGGCTGTCTTGCAAT 82
QY 2402 CATAA 2406
DB 81 CATAA 77
RESULT 9
ADJ38206
ID ADJ38206 standard; DNA; 2283 BP.
XX
XX ADJ38206;
AC
XX
XX 06-MAY-2004 (first entry)
DT
XX
XX Plastid division-related Arc6 orthologue gene 1.
DE
XX
XX prokaryotic type; plastid division; Ftn2; ARC6; ARC5; Fzo; plant cell;
KM agronomic; horticultural; crop plant; ornamental plant; woody plant;
KW herbicide target; gene; ds.
XX
XX Oryza sativa.
OS
XX
XX WO2004001003-A2.
FN
XX
XX 31-DEC-2003.
PD
XX
XX 20-JUN-2003; 2003WO-US019536.
PF
XX
XX 20-JUN-2002; 2002US-0390140P.
PR 09-AUG-2002; 2002US-0402242P.
PR 20-JUN-2003; 2003US-00600070.
XX
XX (UNMS) UNIV MICHIGAN STATE.
PA
XX
XX Osteryoung KW, Vitna S, Koksharova OA, Gao H;
PI WPI; 2004-082486/08.
XX
XX P-PSDB; ADJ38205.
DR
XX
XX New isolated Ftn2, ARC5 and/or Fzo-like nucleic acid sequences, useful
PT for further characterizing plastid division in plant cells, and in
PT varying agronomic and horticultural characteristics of economically
PI important plants.
XX
XX
XX Disclosure; Fig 8; 287pp; English.
PS
XX
XX This invention relates to novel prokaryotic type or plastid division and
CC related genes and proteins. In particular, the invention relates to novel
CC Ftn2 (ARC6), ARC5 and Fzo-like genes and polypeptides. The methods and
CC compositions of the present invention are useful for further
CC characterizing plastid division in plant cells, in order to vary
CC agronomic and horticultural characteristics of economically important
CC plants, such as crop, ornamental and woody plants. They can also be used

CC as herbicide targets. The present sequence is that of a gene which is
CC related to the invention.

XX Sequence 2283 BP; 551 A; 576 C; 592 G; 564 T; 0 U; 0 Other;

Query Match 20.0%; Score 481; DB 12; Length 2283;

Best Local Similarity 54.3%; Pred. No. 6.7e-130;

Matches 1256; Conservative 0; Mismatches 880; Indels 177; Gaps 7;

```
QY 108 CTGCTCGCCGCAAAATGGGCGACGCTTCTCTCCGACCTTCAATTCACTCCGATTC 167
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 132 CTGGGCCCAAGCCCTCTTGCAGATTCACACTCTCCCAAGCCGCGCTTCGACC 191
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 168 CTCCTCTCTCTCTCTGCGCACGCGCACCAACGCACTCTCTCTCTGCAACATC 227
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 192 GCCGTCGCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTT 251
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 228 TATTGATGCTCCGACGCGCACGCTCCCATCCCATTTGATTTCTACAGATTAAGAGC 287
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 252 CCCCCAGCCGCGCAACGCTCCCTCCCAAGTCATTTCTACAGGTTCTAGGGGC 311
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 288 TCAACAATTTCTTAACGATGGAATGGAAGACATTCGAAGCTAGGGTTGAAAC 347
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 312 AGAGCCATTTCTTGGCGATGCAAGAGGCGTTGAGGCAAGATGCGAACCC 371
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 348 GCCGCAATTCGATTTCAGCGACGACGCTTTAATCAGCCGAGACAGATTTCTCAAGCTGC 407
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 372 ACCGAGATGCGCTACAGACGAGATGCTCTTGTGTGTGCGAAGAAATGCTGAGATTGC 431
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 408 TTGCGAAACTGTGTAATCTCGGTCTAGAAAGATGACATGAAGTCTTTTGATGA 467
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 432 CCATGACCTCTCATGAACCAAGAACTCCGCACTCAGATGATGCTGTGCTTTTGAGAA 491
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 468 TGAAGAAGCTACAGTCACTGATGTTCTTGGGTAAGGTTCTGGGGCTCTGTGAT 527
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 492 CCGTGAAGAGCTTCAACATGATATGCTTGGACAAGAGGCTGGG----- 540
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 528 ATTGCAAGAGGTGTGAGACTGAGATAGTTCTTGGGTTGGTGAAGCTCTGCTTAAGA 587
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 541 -----GAGGCACTTGTGCTTGTATCTGAGAAAGAGTTGCTTCTGGA 584
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 588 GAGGTGCTTAAGCTTTTAAGCAAGATGTGTTTATGATGAGCGTTCGTTCTGGA 647
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 585 TCGGCAACCCAGCCCTTCAAGCAAGAGAGTGTGCTACAGATGATTAATGCTGATGGA 644
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 648 TGTCTCGAGGAGTGTATGAGATGAGATCCACCTGATTTTATCTGTTATGATTTGT 707
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 645 TCTATCAAGGAGTGTATGAGCAAGCCCTTCAGATGATTAATGCTGCTCGAGATGCT 704
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 708 TGAGGAAGCTTTGAAGCTTTTACAGAGAGAGAGCAAGTACCTTGCACCGAATTTACG 767
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 705 CGAGAGGCTCTCAAGCTCTTGCAGGAAGATGAGCAAGCAATCTCGACCTGATCTCT 764
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 768 TGCACAAATTGATGAGATCTTGAAGAGATCACTCCGCTTAATGCTTTGAGACTGCG 827
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 765 TTCAAGATTAATGAACTCTCGAGAGATTAACCTGCTGTGTATGAGAGCTTCTCTC 824
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 828 CTACCGCTGTGATGATTAACGTCGAAAGAACTAATATGTTTAACGGGTGCGGAA 887
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 825 CTTTCTATTACACAGACATCAATAAGGCGCAAGAGGCTTCAAGGTGCGAAGAA 884
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 888 TATTTTGTGCTGTGAGAGAGTGAAGATCACTCTTGTGTGGGGTTTGAACCGTGA 947
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 885 CATTTTGTGAGCGTTGCGAGGAGATATTGCTACGTTGAGAGGATTTTCTCGTGA 944
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 948 GAAGTTATGATGAGCGCTTTTACGATGACAGCTCTGAGAGGTTGATCTTTTGT 1007
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 945 AGCTTTCATGAACGAGCTTTTGTGAGATGACATCAATTGAACGATGAGATTTCTTTC 1004
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1008 AGCTACCCCAAGCAATATTCAGAGAGATCAATTGAAGTTTACGAAGTTGAGCTTGCCT 1067
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1005 AAAAACACCAATGACATTTCTCTGAAATGTTGAAATTTTCAATGATGACATTCGACA 1064
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
QY 1068 TGTGCTCAAGCTTTATTTGATAGAACCACTTTTACAGAGATGATTAAGCAATT 1127
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1065 TGTGCTCAAGCAATTAATGATTAAGGCAATTCATCATGATGCGGATGATCTTTT 1124
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1128 CCAGCAACTTCAGAGCTTAAGTAATGCTATGAGAAATTCCTGCATGTTGATGATC 1187
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1125 TGAACCACTCAGAAAGTTCAA-----CATTAAGTCTCAATTA 1160
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1188 ACGGATTAATTTGGAGATAGACTTCGCTAGAAAGGGAGCTGTGACACTGCTTAAGG 1247
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1161 TGCTTATGATTAATGATGATGAGTCTTCATGGAAGGCAATTCCTCATCTGATGCG 1220
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1248 CAAAGTATGATGATGCGGATGATGATGAGCTTACAGATGAGATTCACAAATAGAA 1307
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1221 AGATGATAGCAAGTGCAGAAATGTGCTTGGAAATGATTAATGAGTCTCACATACAGA 1280
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1308 TCCAGCTATTGTGAGATTTTGTGAGAAATTCAAATC---GTATGACATATGATGCT 1364
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1281 CCCCMAAATCTAGAGTTTATGTGACCACTAGCATGAGTGAAGAAATGATCTTCT 1340
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1365 CCTGACATATGCAAAATTTGTGAAACCTGATGAGAGGGTGTCTTCTAGATTCAAG 1424
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1341 TCCAGGCTGTGCAAGCTTTTGGAGACTTGTATCTTGAAGTTTCTTAGAGACAG 1400
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1425 AGACACCAAGATTAATAAATTTAACTGGGAGCTACTATGATGATCTATGTTTGA 1484
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1401 AGATATCGGGGCAATGCAATTCAGACTTGGAGATTACTAGATGATCAGAAATTTAAG 1460
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1485 TTACTTGGAAAGATGAGATGATTCAGAGGTTCTCCCTTATGCTGCTGCACTAATGC 1544
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1461 CTACTTGAAGAGATGAGAGGTGTGTGCTCTCATTTGAGCTGCTGCTGATATGC 1520
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1545 AAGGATGAGACCGAGCATGTGAAGCTAGTCTATGACAGCACTGCAAGAAATTTTCC 1604
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1521 AAAACTTGTGTCTACACTACAGCTGCACTTGGTA----- 1555
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1605 TTCCGCTATACAGATGAATACTCGCTGAACCCAGAGATGTGCAGAGACAGTATTAG 1664
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1556 -----CTGGAATCAAAATGATTAACAAG----- 1581
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1665 TGTATATCTGTGTGTAACAATGATGAGCGGTGATGTGAGACCTGTGTCTTATGAGA 1724
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1582 -----TTCAACAA 1589
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1725 AGCTGAAGACCCCTGAAAACCTTGAACCTAATGATTAATGCAATTGAGCTGGGGCTC 1784
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1590 GGTTTTCCATTGATAGAACATGTTAGACAGTCAAGCATGAAATCTAAGATGCGCC 1649
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1785 AGAGATGAGGTGATGAACCTACTGTGAAATGTCCGTTGCTGATATGTTAAGAGGC 1844
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1650 TGGGAGATATCTGAAATTTTGAACAGAAATGACCTGCTCATGATTCGAGAAATGC 1709
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1845 AAGTGTGAATCTTACCTGTGTGTGTGAGCAATTGACATTAATCACTGTTCAAGCCGAA 1904
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1710 CGCCTTAAGATTAATCTCTGTGCGCATGTTTGCATGTTGCAATTAATGGGGCCAA 1769
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1905 GATTTTCTTAAAAGCACTCATCTTTTCAACCAAGAGATAGTATTTCTTATGATATC 1964
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1770 ATATTTGCTCTGTAAGAGGCGCCCTTCTGTCTATTAAGAGTGAAGATGATCTGTGGAGT 1829
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1965 TGAT-----GTGCTACATAGGGTCACTGAGAGCTGACGATTCAGAGCACTTCC 2015
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2016 CAGAAATGATGCTAGACTGACAGAAATATAGATATCAATGCGACAGAAATTAAGTCT 2075
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1990 TAGAATGATGCGAAGCTGCGAAGAAATATGTTGCGAAATGCGAGATCAATCTTAA 1949
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2076 GGCCTTTGGGCTGATACCGCATAGAAATGTTTACAGAGATTTTGGATGGGCAATGCT 2135
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1950 GGCCTTGGACCAAGACATTCGTTGCAATTTGCAAGAGATTTCTTGAATGCAACATGCT 2009
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2136 GAAGATTTGACATGACAGACAGCTGAACCTGCGACGCTTGGGTTGTTATGATTAATAC 2195
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
Db 2010 AAGGCTGACCTGACCGAGCGAGATTGAGCGCTCAGTGGGTGTTCTGGAGATATAC 2069
Oy 2196 ACTGTTGAACATATCTGTTGACAGTGTGACAGTCTCAGACAGATGAAACCGTGTCTGT 2255
Db 2070 ACTATCCAGATGTGACAGATTGATGATCATCTATCTCCATGATGTGACGAGCACTGT 2129
Oy 2256 GGAAGCACTGTGAGAGCTGTGTTCTATCTGATTTGTTGTTCAATCCAGAAAACATGC 2315
Db 2130 GGAGGCTACGATTGATGAGGCGAGCCCACTTACTGATGTACTGAGAGCCAGAAACATGA 2189
Oy 2316 TACTGATGTGACAGATCTACACAAAGATACGAAGTTTCTGTCCAG---TGAGGCTG 2372
Db 2190 TTCATATGACACAAATATACATACCGGTATGATGAGTGGCTTCTCCAACTAGAGGGTG 2249
Oy 2373 GAAATCACTGAAGGCTGTGTTCTTGATCATA 2405
Db 2250 GAGATACGGAAGGAGCAGTCTCTCAAGTCGTA 2282

RESULT 10
ADJ38264
ID ADJ38264 standard; cDNA; 631 BP.
XX
AC ADJ38264;
XX
DE 06-MAY-2004 (first entry)
XX
DE Plastid division-related Arc6 orthologue cDNA 51.
XX
KM prokaryotic type; plastid division; Ftn2; ARC6; ARCS; Fzo; plant cell;
KM agronomic; horticultural; crop plant; ornamental plant; woody plant;
KM herbicide target; gene; ss.
XX
OS Prunus persica.
XX
PN WO2004001003-A2.
XX
PD 31-DEC-2003.
XX
PF 20-JUN-2003; 2003WO-US019536.
XX
PR 20-JUN-2002; 2002US-0390140P.
XX
PR 09-AUG-2002; 2002US-0402242P.
XX
PR 20-JUN-2003; 2003US-00600070.
XX
PA (UNMS ) UNIV MICHIGAN STATE.
XX
PI Oosteryoung KW, Vitha S, Koksharova OA, Gao H,
PI WPI; 2004-082486/08.
XX
DR WPI; 2004-082486/08.
XX
PT New isolated Ftn2, ARCS and/or Fzo-like nucleic acid sequences, useful
PT for further characterizing plastid division in plant cells, and in
PT varying agronomic and horticultural characteristics of economically
PT important plants.
XX
PS Disclosure; Fig 8; 287pp; English.
XX
CC This invention relates to novel prokaryotic type or plastid division and
CC related genes and proteins. In particular, the invention relates to novel
CC Ftn2 (ARC6), ARCS and Fzo-like genes and polypeptides. The methods and
CC compositions of the present invention are useful for further
CC characterizing plastid division in plant cells, in order to vary
CC agronomic and horticultural characteristics of economically important
CC plants, such as crop, ornamental and woody plants. They can also be used
CC as herbicide targets. The present sequence is a cDNA sequence which is
CC related to the invention.
XX
SQ Sequence 631 BP; 174 A; 123 C; 155 G; 178 T; 0 U; 1 Other;
Query Match 13.7%; Score 328.8; DB 12; Length 631;
Best Local Similarity 71.5%; Pred. No. 1.8e-85;
```

```
Matches 451; Conservative 0; Mismatches 168; Indels 12; Gaps 1;
Oy 913 GGAGCATCAGCTCTGTTGGGGGTTTGAACCCGTGAGATTATGATAGAGCGTTT 972
Db 1 GCGATTGCAATTTCTGGGGGNGATTCACTAGTGAAAATTTTCATBAAGAGGCTTCTTG 60
Oy 973 CGAATGACAGCTGCTGAGCAGGTGATCTTTTGTAGCTACCCCAAGCAATATTTCCAGCA 1032
Db 61 CATATGACTGACGCTGACAGGTTGATTTATTTGAGTACCCCAATATTTCCGCGCA 120
Oy 1093 GAGTCATTTTGAAGTTTAAAGTTGCACTGCTCTGTGGGCTCAAGCTTTTATTTGTTAG 1092
Db 121 GAAAGCTTGAAGTTTATGGGGTGGCTTGTGCGCTGTGTGCTCAAGCTTTGTGTGTA 180
Oy 1093 AAGCCACACCTTTTACAGAGATGCTGATTAAGCAATTTCCAGAACTTACAGAGCTTAAGSTA 1152
Db 181 AAACCTCATCATTTCAAGATGCTGAAAACCTATTTCCAGAACTTACAGAGCTTAAGSTA 240
Oy 1153 ATGGCTATGAGAGATTCTTCGATGTTGATGATACCGAAATTAATTGGAGATAGACTTC 1212
Db 241 ACAGCTGTAGGACATTTCTTGTGACAACTATATTAACAAAGAAAGAGATGAGATAGACTTT 300
Oy 1213 GGTCTAAGAAAGGGACCTGTGCACTGCTTATAGCCAAAGTTGATGATGCCGTATGCG 1272
Db 301 GCTTTGAGAGGGGACCTCTGTTCACTTCTTACGGGACCTTGATGACAGTCGTTCCGCG 360
Oy 1273 TTGGGCTTAGACAGTGAAGATTCACAATATAGGAATCCAGCTAATTTGAGATTGTTTG 1332
Db 361 TTGGGCTTAGACAGTGAAGATTCACAATATAGGAATCCAGCTAATTTGAGATTGTTTG 420
Oy 1333 GAGAATTCAAATTCGTGATG-----ACAATGATGATCTCCCTGACATATGCAAA 1380
Db 421 GAGAACTCAAGATGACGATGACATGACATGACATGATGATCTCCCTGACATTTGCAAG 480
Oy 1381 TTGTTGAAACCTGTTGGGCGAGGGTGTCTTTCTTACGTTCAAGAGACCAAGATATAA 1440
Db 481 CTATTTGAGAGCTGTGTATGATGAGGTGTATTTCCAGGTTTAAAGACCAAGAGACATA 540
Oy 1441 AAATTTAACTCGGGGACTACTATGATGATCTCTATGTTTGAAGTTACTTGGAAAAGAGTG 1500
Db 541 GAGTTCAAGCTGGAGACTACTATGATGATCTCTACAGTCTTGAGATACTTGGAAAAGGCTG 600
Oy 1501 GAGGTAGTTCAAGGTTCTCTTACGTCGCTG 1531
Db 601 GATGGCACTAATGTTGACCTTAGCTGCTG 631

RESULT 11
ADJ38215
ID ADJ38215 standard; cDNA; 660 BP.
XX
AC ADJ38215;
XX
DE 06-MAY-2004 (first entry)
XX
DE Plastid division-related Arc6 orthologue cDNA 8.
XX
KM prokaryotic type; plastid division; Ftn2; ARC6; ARCS; Fzo; plant cell;
KM agronomic; horticultural; crop plant; ornamental plant; woody plant;
KM herbicide target; gene; ss.
XX
OS Medicago truncatula.
XX
PN WO2004001003-A2.
XX
PD 31-DEC-2003.
XX
PF 20-JUN-2003; 2003WO-US019536.
XX
PR 20-JUN-2002; 2002US-0390140P.
XX
PR 09-AUG-2002; 2002US-0402242P.
XX
PR 20-JUN-2003; 2003US-00600070.
```

PA (UNMS) UNIV MICHIGAN STATE.
XX
PI Oosteryoung KW, Vitha S, Koksharova OA, Gao H;
XX MPI; 2004-082486/08.
DR
XX
PT New isolated Ftn2, ARCS and/or Fzo-like nucleic acid sequences, useful
PT for further characterizing plastid division in plant cells, and in
PT varying agronomic and horticultural characteristics of economically
PT important plants.
XX
PS Disclosure: Fig 8; 287pp; English.
XX
XX This invention relates to novel prokaryotic type or plastid division and
XX related genes and proteins. In particular, the invention relates to novel
XX Ftn2 (ARCS), ARCS and Fzo-like genes and polypeptides. The methods and
XX compositions of the present invention are useful for further
XX characterising plastid division in plant cells, in order to vary
XX agronomic and horticultural characteristics of economically important
XX plants, such as crop, ornamental and woody plants. They can also be used
XX as herbicide targets. The present sequence is a cDNA sequence which is
XX related to the invention.
SQ Sequence 660 BP; 149 A; 178 C; 152 G; 181 T; 0 U; 0 Other;
Query Match 11.6%; Score 278.8; DB 12; Length 660;
Best Local Similarity 67.3%; Pred. No. 9.6e-71;
Matches 432; Conservative 0; Mismatches 192; Indels 18; Gaps 2;
QY 81 ACCGAGCCACAACACCTTACACATATGCTCCGCCAGCAAAATGGGCGGACCGTCTT 140
DB 17 ACCTAACCGTCTCAATTCCTCCGCGTCCGCAACGATTAATGGGCGGACCACTCAT 76
QY 141 CTCGACTTCATTTACCTCCGATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 200
DB 77 TTCGATTTCAATTCCTCGGCGACACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 136
QY 201 CGCCACTCTGCTCTCTGCGACCATCTAATGATGTCGCCAAGCGCAGTCCCATCC 260
DB 137 AGTCACTCTC-----ACTCTCTTACCTCTCTCCGATAGAGCGCCAGTCACTCC 190
QY 261 CATGATTTTACGAGATTAAGAGCTCAACATCTTTTACCCAGTGAATCAGAG 320
DB 191 TCTGACCTGTACAAATCTCTCGGCGCAAAACGATTTCTCGGTGATTCGAG 250
QY 321 AGCATTCGAACTAGGGTTTGGAAACCGCGCAATTCGGTTTCAAGCAGCGTTTAA 380
DB 251 AGCTTATGAAGCAATCTGGAAGCTCTCAAGATCTTCAATATGAAGCTTCAT 310
QY 381 CAGCCGAGACAGATTCTTCAAGCTGCTTGGAAACCTGTCTAATCTCGGTCTAGAG 440
DB 311 TAGTCGTGTGAGATTCTTCAAGCTGCTTGGAAACCTAGCTGATCTCTCTAGAG 370
QY 441 AGACTACAAATGAAGTCTTCTGATGATGAA-----GAGTACAGTCATCAG 488
DB 371 AGAGTATATGAAGCTCTGTGAGATGAAGACGAGATGAGATTTTCATTTCTAC 430
QY 489 TGATGTCCTTGGGATAGGTTCTCGGGGCTCTGTGATTTGCAAGAGGTGGAGAG 548
DB 431 TGAATTCCTTTGCAAAAGTTCTGAGCTCTGTGCTGTGGCAAGAGCTGAGAGAG 490
QY 549 TGAATAGTTCTTGGGTTGTGAGGCTCTGCTTAAAGAGAGGTTCCCTAAGTGTAA 608
DB 491 GGAATGTGCTCGGATTTGAGGGGGTTTACTGAGAGAGAGGTTTACGAGAGATGTTAA 550
QY 609 GCAAGATGTGTTTAAATAGTATGAGGCTGCTTCTGATGCTGAGAGATGCTATGAG 668
DB 551 GCAAGATGTGTGTGCTATGAGGCTTGAATGTGACGTTCTTCAAGAGATGCTATGAG 610
QY 669 ATTGATCCACCTGATTTTATTAATGATGATGATTTGTTGA 710
DB 611 TTGTGCCCCGCAATTTTCAATTTGTTGCTGTGAGATGCTGGA 652

RESULT 12
ADJ38223
ID ADJ38223 standard; cDNA; 537 BP.
XX
XX
AC ADJ38223;
XX
XX 06-MAY-2004 (first entry)
DT
XX
XX Plastid division-related Arc6 orthologue cDNA 16.
DE
XX
XX prokaryotic type; plastid division; Ftn2; ARCS; Fzo; plant cell;
XX agronomic; horticultural; crop plant; ornamental plant; woody plant;
XX herbicide target; gene; ss.
OS
XX Triticum aestivum.
PN
XX MO2004001003-A2.
PD
XX
XX 31-DEC-2003.
PE
XX 20-JUN-2003; 2003WO-US019536.
PR
XX 20-JUN-2002; 2002US-0390140P.
PR 09-AUG-2002; 2002US-0402242P.
PR 20-JUN-2003; 2003US-00600070.
XX
XX (UNMS) UNIV MICHIGAN STATE.
PA
PI Oosteryoung KW, Vitha S, Koksharova OA, Gao H;
XX MPI; 2004-082486/08.
DR
XX
XX New isolated Ftn2, ARCS and/or Fzo-like nucleic acid sequences, useful
PT for further characterizing plastid division in plant cells, and in
PT varying agronomic and horticultural characteristics of economically
PT important plants.
XX
XX
XX Disclosure: Fig 8; 287pp; English.
PS
XX
XX This invention relates to novel prokaryotic type or plastid division and
XX related genes and proteins. In particular, the invention relates to novel
XX Ftn2 (ARCS), ARCS and Fzo-like genes and polypeptides. The methods and
XX compositions of the present invention are useful for further
XX characterising plastid division in plant cells, in order to vary
XX agronomic and horticultural characteristics of economically important
XX plants, such as crop, ornamental and woody plants. They can also be used
XX as herbicide targets. The present sequence is a cDNA sequence which is
XX related to the invention.
SQ Sequence 537 BP; 133 A; 113 C; 151 G; 140 T; 0 U; 0 Other;
Query Match 9.3%; Score 224; DB 12; Length 537;
Best Local Similarity 63.6%; Pred. No. 1.1e-54;
Matches 341; Conservative 0; Mismatches 195; Indels 0; Gaps 0;
QY 555 AGTTCCTGGGTTGTTGAGGCTCTGCTTAAAGAGAGTTGCTTAAGCTTTAAGCAAGA 614
DB 2 AGTGTCTGCAATTTGAGGGGCACTTACGAGAGACCGCCGCCAAGGCTTCAAGCAAGA 61
QY 615 TGTGTTTAAATAGGCGCTTGGCTTCTGATGTCTGAGGAGATGATGCAATTGGA 674
DB 62 TGTGTTGCTGGCAATGCGCTCTTATGTGATCTATCAAGGAGAGCAATGCGGCTAG 121
QY 675 TCAACCTGATTTTATTAATGCTTATAGATTGTTGAGGAAGCTTTGAAGCTTTACAGA 734
DB 122 CCGTCAAGATGTAATCCGCTGCTGTGAGTGCTTGAAGGCTCTCAAGCTTTTGACAGA 181
QY 735 GGAAGAGCAATGAGCTTCTCACCGATTTACGTCGACCAATTTGATGAGACTTTGAGAGA 794
DB 182 GGAATGGGCAATCAATCTGACACTGGTTGCTCTCAAAATTTGATGAGAACTCTGAGAGA 241
QY 795 GATCACTCCGCGTTATGTCTTGAAGCTACTTGGCTTACCGCTTGGTATGATTAACGCTGC 854

Db 242 TATCACACCTCTGTGTGTTTGAGACCTTCTGCCCTTCTCTGTATGAAAAACATCAGAA 301
QY 885 GAAAAGACCTAATGGTTTAAGCGGTGTGGGAAATTTTGTGTCGTTGGAGAGTGGA 914
Db 302 TGAACACCAAGAGGTCTTCGTGTGTGAGAAACATTTTGTGATGTTGGACAGAGAG 361
QY 915 AGCATCAGCTCTTGTGTGGGGTGTGAACCGGTGAGAAAGTTATGAATGAGCCGTTTACG 974
Db 362 TATTGTAATCTGTGTGAGAGGATTTTCGGTGAAGCCTACATGAATGAAGCCTTCTGCA 421
QY 975 AATGACAGTGTGTGACGAGTTGATCTTTTGTAGTACCCACACATATTCAGAGAGA 1034
Db 422 GATGACATCGCGGAGAGATGATGATTTCTTCCAAAACACCGAATAGCATCCCTCGA 481
QY 1035 GTCATTTGAAGTTTACGAAGTTGACCTGTGTGGCTCAAGCTTTTATTTGTA 1090
Db 482 ATGTTTGAATCTTATAGCGTGGACCTTGCAAAATGTGCTCAAGCAATTTGAATGA 537

RESULT 13

ACN48855/C
ID ACN48855 standard; cDNA; 552 BP.

AC ACN48855;

DT 02-DEC-2004 (first entry)

DE Cotton primed seed EST Clone ID: LIB3825-027-Q6-N6-H1, SEQ:3636.

XX Cotton; plant; EST; expressed sequence tag; transgenic plant; seed;
KM variety DP50B; library LIB3825; molecular tag; molecular marker;
KM genetic mapping; molecular mapping; seed germination; plant growth;
KM plant quality; plant yield; plant breeding; tissue printing; ss.

OS Gossypium hirsutum.

XX US2004123340-A1.

XX PD 24-JUN-2004.

XX PF 12-DEC-2001; 2001US-00021323.

XX PR 14-DEC-2000; 2000US-0255619P.

XX PA (DEIK/) DEIKMAN J.

XX PA (FENG/) FENG P C C.

XX PA (FINC/) FINCHER K L.

XX PA (ZIEG/) ZIEGLER T E.

XX PI Deikman J, Feng PCC, Fincher KL, Ziegler TE;

XX WPI; 2004-479808/45.

XX New isolated nucleic acid molecule that encodes a plant protein or its
PT fragment, useful for isolating a variety of agronomically significant
PT genes associated with plant growth, quality or yield, and as molecular
PT tags to map genes.

XX Claim 1; SEQ ID NO 3636; 34pp; English.

XX The invention relates to 17880 cotton expressed sequence tags (ESTs;
CC ACN45220-ACN63099). The ESTs were isolated from cDNA libraries generated
CC from primed or non-primed seeds from variety DP50B, mature seeds from
CC variety Coker 312 Boswell 96 Field, and androecium tissue, gynoecium
CC tissue, developing fibres, carpel walls and sepal from variety
CC Nucleon338. The invention also relates to substantially purified
CC proteins or their fragments encoded by nucleic acid molecules of the
CC invention, and to transformed plants having a nucleic acid construct
CC comprising a nucleic acid of the invention. The cotton ESTs are useful as
CC molecular tags to isolate genetic regions, to isolate genes, to map
CC genes, to determine gene function and to determine whether genes are
CC members of a particular gene family. The nucleic acid molecules may be

CC used for isolating a variety of agronomically significant genes
CC associated with plant growth, quality, yield, and could also serve as
CC links in metabolic and catabolic pathways. The nucleic acid molecules are
CC also useful for identifying genes important in initiating and maintaining
CC seed germination or that may be used to mitigate stresses encountered
CC during seed germination. The ESTs additionally enable the acquisition of
CC promoters and cis-regulatory elements which will be useful to express
CC agronomically significant genes in these tissues and/or other tissues,
CC and also permits the acquisition of molecular markers useful in breeding
CC schemes, genetic and molecular mapping, and in cloning of agronomically
CC significant genes. The nucleic acid molecules are further useful for
CC detecting the expression level or pattern of a protein or mRNA and for
CC detecting the presence or quantity of a protein by tissue printing. The
CC present sequence represents a specifically claimed EST isolated from a
CC cotton variety DP50B primed seed cDNA library (LIB3825). The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from the US patent office at
CC seqdata.uspto.gov/sequence.html?docID=US20040123340
XX

SQ Sequence 552 BP; 153 A; 122 C; 110 G; 167 T; 0 U; 0 Other;

Query Match 9.2%; Score 222.4; DB 13; Length 552;

Best Local Similarity 72.2%; Pred. No. 3.4e-54;

Matches 289; Conservative 0; Mismatches 111; Indels 0; Gaps 0;

QY 2006 AAGCACTTCCAGATGAGTCTAGAGCTGCGAGAAATATAGTATCCAGTGGCAGAGA 2065

Db 550 AAGAATTACTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 491

QY 2066 TTAAGTCTCTGCTTTGGGCTGATACCGCATAGAAATGTACCAAGAGTTTGGATG 2125

Db 490 TTAATATGAGGCGTTTGGACCTTATACCCCTTGATTAATTTGCCAAGAGTTCTGATG 431

QY 2126 GGCGAATGCTGAAGATTGACTGACAGAGCAGCTGAAACTGCGAGCTTGGGTTGTTT 2185

Db 430 GTCAATTTTGAACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 371

QY 2186 ATGATTTATACCTGTGAACTATCTGTGACAGTGTGACAGTCTCAGATGAGACC 2245

Db 370 ATGATATATAGCTACTGAAACATGACATGACATGATGATGATGATGATGATGATG 311

QY 2246 GTGCTGTGTTGAAAGCACTTGGAGAGTGTGCTGTATCTGATTTGGTTCAATCC 2305

Db 310 GAGCTGTATGTTGAAGCTACTCTGGAAGATTCACCTGTCTTGAATGATTCATCCGG 251

QY 2306 AAAACAATGCTACTGATGTTGAACTTGAACCTTATCTGTTGACAGTCTCAGATG 2365

Db 250 AGAACATATGCTCTTAATGTAACTCTTACACACGAGATATGATGTCTTGTTCAACT 191

QY 2366 CAGGTTGAAATCACTGAAAGCTCTGTTCTTGATCATTA 2405

Db 190 CAGCTGGAATATCACTGAAGATCTGTCTTCAATCTTA 151

RESULT 14
ADJ38255
ID ADJ38255 standard; cDNA; 545 BP.

AC ADJ38255;

DT 06-MAY-2004 (first entry)

DE Plasmid division-related Arc6 orthologue cDNA 42.

XX prokaryotic type; plasmid division; Ftnr; ARC6; ARC5; Fzo; plant cell;

XX agronomic; horticultural; crop plant; ornamental plant; woody plant;

XX herbicide target; gene; ss.

OS Gossypium arboreum.

XX WO200401003-A2.

XX 31-DEC-2003.

Db	362	CAATCAACCTTGACCTGCTCTGCTTTCAAAATTGATGAAGAACTCTGGAGAGATCACC	421
Qy	803	CGCGTTATGCTCTTGGAGCTACTTGGCTTACCGCTTGATGA	842
Db	422	CTCGTTGCTTTTGGAGCTTCTTGCCCTTNCCTTGATGA	461

Search completed: December 10, 2005, 19:21:37
Job time : 1376.27 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 10, 2005, 16:58:05 / Search time 9151.75 Seconds
(without alignments)
12300.344 Million cell updates/sec

Title: US-10-600-070B-1
Perfect score: 2406
Sequence: 1 atggaagctcgtacgtacgt.....gtctctctctgcacataa 2406

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues
Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST: *
1: gb_eac1: *
2: gb_est2: *
3: gb_est3: *
4: gb_hic: *
5: gb_est4: *
6: gb_est5: *
7: gb_est6: *
8: gb_est7: *
9: gb_gsa1: *
10: gb_gsa2: *
11: gb_gsa3: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	902.2	37.5	919	4	CNS09YXH
2	633	26.3	741	4	CNS09YXH
3	633	26.3	741	4	CNS09YXH
4	482	20.0	561	1	AI998415
5	458	19.0	534	1	BO834167
6	447	18.6	2307	10	CL965374
7	436	18.1	813	7	CO071968
8	429.4	17.8	897	6	CD573714
9	360.8	15.0	785	9	BZ437564
10	359.2	14.9	624	8	CO075595
11	344.2	13.3	769	7	CO075595
12	335.4	13.9	741	8	CO075595
13	328.8	13.7	631	7	CO075595
14	326.6	13.6	703	7	CNS18842
15	321	13.3	415	3	BP785511
16	316	12.1	410	3	BP785511
17	306.8	12.8	611	8	BP782886
18	304.4	12.7	341	9	AY199896
19	300.6	12.5	728	8	AY199896
20	299.2	12.4	434	3	BP620404
21	293.4	12.2	832	7	CO079829
22	288.2	12.0	697	8	CX173275

C 23	287.6	12.0	821	7	CK090561	CK090561
C 24	287.6	12.0	922	7	CV264112	CV264112
C 25	286.6	11.9	819	7	CV241483	CV241483
C 26	283	11.8	898	8	DR830109	DR830109
C 27	280.2	11.6	722	7	CO117046	CO117046
C 28	278.8	11.6	660	2	B1268376	B1268376
C 29	272.2	11.3	839	7	CK936162	CK936162
C 30	270.4	11.2	866	6	CD573715	CD573715
C 31	270.2	11.2	723	7	CO079828	CO079828
C 32	265	11.0	746	8	CK674369	CK674369
C 33	261.8	10.9	768	8	CK543654	CK543654
C 34	260.2	10.8	594	1	AM696905	AM696905
C 35	257.4	10.7	573	7	CNS904734	CNS904734
C 36	254	10.6	849	7	CO117047	CO117047
C 37	249.6	10.4	871	10	DU050213	DU050213
C 38	235	9.8	820	8	DR972351	DR972351
C 39	231.8	9.6	751	11	CR486014	CR486014
C 40	226.6	9.4	521	7	CNS14655	CNS14655
C 41	226	9.4	683	7	CO076192	CO076192
C 42	224	9.3	537	2	BE490117	BE490117
C 43	220	9.1	501	5	BU049629	BU049629
C 44	218.6	9.1	508	3	BU048486	BU048486
C 45	214.8	8.9	451	1	AJ802091	AJ802091

ALIGNMENTS

RESULT 1
CNS09YXH 919 bp mRNA linear HTC 04-FEB-2004
LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone
DEFINITION GSUTS89ZC08 of Adult vegetative tissue of strain col-0 of
Arabidopsis thaliana (thale cress).

ACCESSION BX841670.1 GI:42406830
VERSION BX841670.1
KEYWORDS HTCC; GSUT cDNA.
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana (thale cress)

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eustroide II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 919)
Castelli, V., Aury, J.M., Jallion, O., Wincker, P., Clepet, C.,
Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V.,
Temple, G., Caboche, M., Weissenbach, J. and Salanoubat, M.
Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation

TITLE Arabidopsis thaliana Full-length cDNA Complete sequence from clone
AUTHORS Castelli, V., Aury, J.M., Jallion, O., Wincker, P., Clepet, C.,
Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V.,
Temple, G., Caboche, M., Weissenbach, J. and Salanoubat, M.

JOURNAL Unpublished
REFERENCE 2 (bases 1 to 919)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

COMMENT The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences). 5 prime and 3 prime are assembled with Ptipap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
length

FEATURES
source Location/Qualifiers
1..919
/organism="Arabidopsis thaliana"
/mol_type="mRNA"

http://www.genoscope.cns.fr/cgi-bin/gsb/gsb?source=Arabidopsis.
Location/Qualifiers
1..919
/organism="Arabidopsis thaliana"
/mol_type="mRNA"

/db_xref="taxon:3702"
/clone="GSITLS92C08"
/tissue_type="adult vegetative tissue"
/ecotype="Col-0"
/plasmid="pCMVSPORT_6"
1..919
/gene="At5g42480"

Query Match 37.5%; Score 902.2; DB 4; Length 919;
Best Local Similarity 99.7%; Pred. No. 4,7e-244;
Matches 904; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

1 ATGGAAGCTCTGAGTACAGTCGCAATGGTCTCTCCCAATTCATTAATGCCGATTACCA 60
13 ATGGAAGCTCTGAGTACAGTCGCAATGGTCTCTCCCAATTCATTAATGCCGATTACCA 72
61 CCGGAGAGCAAAAGCTCGAGAGTGGCCCAACACCTTCAACAATATCTGCTCCGCGAC 120
73 CCGGAGAGCAAAAGCTCGAGAGTGGCCCAACACCTTCAACAATATCTGCTCCGCGAC 132
121 AAATGGGCGAGCGCTTCTCTCGACTTCAATTCACCTCGATTCCTCCCTCTCTCC 180
133 AAATGGGCGAGCGCTTCTCTCGACTTCAATTCACCTCGATTCCTCCCTCTCTCC 192
181 TTGCGCACCGCCACCAACCGCCACTCTCGTCTCTGCGACCACTTAATGATGATGCC 240
193 TTGCGCACCGCCACCAACCGCCACTCTCGTCTCTGCGACCACTTAATGATGATGCC 252
241 GAAGCCGACGTCCTCCATCCCATTTTATTTTCAACGATTTTGAAGGCTCAAAACATTT 300
253 GAAGCCGACGTCCTCCATCCCATTTTATTTTCAACGATTTTGAAGGCTCAAAACATTT 312
301 TTAACCGATGATGAGAGAGATTCGAAAGTGAAGGTTTGAAGGCTTGAAGGCTTGAAG 360
313 TTAACCGATGATGAGAGAGATTCGAAAGTGAAGGTTTGAAGGCTTGAAGGCTTGAAG 372
361 TTACGCGAGAGAGCTTTTATATACGCGGAGACAGATTTCTTCAAGCTGCTTGGCAAA 420
373 TTACGCGAGAGAGCTTTTATATACGCGGAGACAGATTTCTTCAAGCTGCTTGGCAAA 432
421 TCTATCTCTCGGTCTTGAAGAGAGATTCGAAAGTGAAGGCTTCTTGAATGAGAGAGAT 480
433 TCTATCTCTCGGTCTTGAAGAGAGATTCGAAAGTGAAGGCTTCTTGAATGAGAGAGAT 492
481 GTATCATCATGATGTCCTTGGGATTAAGGTTCTGGGGCTCTCTGTGATTTGCAAGAGAT 540
493 GTATCATCATGATGTCCTTGGGATTAAGGTTCTGGGGCTCTCTGTGATTTGCAAGAGAT 552
541 GGTGAGACTGAGATTAAGTCTTGGGTTGGTGAAGGCTTCTTGAAGAGAGGTTGCTTAAG 600
553 GGTGAGACTGAGATTAAGTCTTGGGTTGGTGAAGGCTTCTTGAAGAGAGGTTGCTTAAG 612
601 TCGTTTAAGCAAGAGTGGTTTAAATGATTAAGGCTTGGGCTTCTGATGTCGAGAGAT 660
613 TCGTTTAAGCAAGAGTGGTTTAAATGATTAAGGCTTGGGCTTCTGATGTCGAGAGAT 672
661 GCTATGAGCTGAGATTCACCTGATTTTATTAATCTGTTAAGTTTGAAGAGAGCTTTG 720
673 GCTATGAGCTGAGATTCACCTGATTTTATTAATCTGTTAAGTTTGAAGAGAGCTTTG 732
721 AAGCTTTTAAGAT 780
733 AAGCTTTTAAGAT 792
781 GAGACTTTGAAGAGAGATCACTCCGCTTATGCTTGAAGAGATCACTCCGCTTATGCT 840
793 GAGACTTTGAAGAGAGATCACTCCGCTTATGCTTGAAGAGATCACTCCGCTTATGCT 852
841 GATGATTAAGAT 900
853 GATGATTAAGAT 912

QY 901 GTTGAGAG 907
DB 913 GTTGAGAG 919

RESULT 2

CNS09YNM 741 bp mRNA linear HTC 06-FEB-2004
LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone
DEFINITION GSITLS92A05 of Silique of strain col-0 of Arabidopsis thaliana
(thale cress).

ACCESSION BX33489.1 GI:42455179
VERSION BX33489.1
KEYWORDS HTC; GSIT_cDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana

REFERENCE Castelli,V., Aury,J.M., Jallion,O., Wincker,P., Clepet,C.,
Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,
Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.,
Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation

TITLE Arabidopsis thaliana
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 741)
GENOSCOPE Genoscope.
AUTHORS Direct Submission
TITLE Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seget@genoscope.cns.fr
JOURNAL

COMMENT The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences) . 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_projet_EF/Full
length
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.

FEATURES
source location/Qualifiers
1..741
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="GSITLS92A05"
/tissue_type="Silique"
/ecotype="Col-0"
/plasmid="pCMVSPORT_6"
1..741
/gene="At5g42480"

ORIGIN
Query Match 26.3%; Score 633; DB 4; Length 741;
Best Local Similarity 100.0%; Pred. No. 1.1e-167;
Matches 633; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1774 GCTGGGCTCAGAT 1833
1 GCTGGGCTCAGAT 60

DB 1774 TTAAGAT 1893
1774 TTAAGAT 120

QY 1894 TTGAGCGAAGATTTTCTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1953
1894 TTGAGCGAAGATTTTCTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1953

Db		121	TTCACCCAGAAAGTATTTCTTAAAGCAGCTCATCTTTTCAACGGAAGATATGGTTCT	180
Oy		1954	TCTATGGAATCTGATGTGGCTACTCATTAGGCTAATGCAGAGCTGACGATTCAGAGACATT	
Db		161	TCTATGGAATCTGATGTGGCTACTCATTAGGCTAATGCAGAGCTGACGATTCAGAGACATT	240
Oy		2014	CCCGAATAGSAGTGTAGGACTGCAGAGAATATAGATCCAAGTGCCAGAAATTAAGTCT	2073
Db		241	CCCAGAAATGATGCTTAGAGACTGCAGAGAATATAGATCCAAAGTGCCAGAAATTAAGTCT	300
Oy		2074	CTGGCTTTTGGGCCCTGATCACCCGATAGAAAATGTTACAGAGTTTTGGATGGCGAATG	2133
Db		301	CTGGCTTTTGGGCCCTGATCACCCGATAGAAAATGTTACAGAGTTTTGGATGGCGAATG	360
Oy		2134	CTGAAGATTTTGGACTGACAGAGCAGCTGAACTCGCAGCTTTGGTGGTTATGATTAAT	2193
Db		361	CTGAAGATTTTGGACTGACAGAGCAGCTGAACTCGCAGCTTTGGTGGTTATGATTAAT	420
Oy		2194	AACATGTTGAACATACTGTTGACAGTGTGACAGCTCTAGACAGATGAAACCCCGTCTTG	2253
Db		421	AACATGTTGAACATACTGTTGACAGTGTGACAGCTCTAGACAGATGAAACCCCGTCTTG	480
Oy		2254	GTCGAAGCAACTCTGAGAGAGTCTGCTGTCTATGTGATTTGGTTCATCCAGAAAACAAT	2313
Db		481	GTCGAAGCAACTCTGAGAGAGTCTGCTGTCTATGTGATTTGGTTCATCCAGAAAACAAT	540
Oy		2314	GCTACTGATGTGAGAACCTTACACAACAAGATACGAAGTTTTCTGCTCCAACTCAGGCTGG	2373
Db		541	GCTACTGATGTGAGAACCTTACACAACAAGATACGAAGTTTTCTGCTCCAACTCAGGCTGG	600
Oy		2374	AAATCACTGAAGGCTCTGTTCTTGATCATATA	2406
Db		601	AAATCACTGAAGGCTCTGTTCTTGATCATATA	633
RESULT 3 CNS09Z55				
LOCUS DEFINITION	CNS09Z55	741 bp	mRNA . linear	HTC 06-FEB-2004
			Arabidopsis thaliana Full-length cDNA Complete sequence from clone GSIJTSII312E09 of Silique of strain col-0 of Arabidopsis thaliana (thale cress) .	
ACCESSION VERSION	BX833051	GI:42455503		
KEYWORDS SOURCE	HTC; GSRLT_cDNA.			
ORGANISM	Arabidopsis thaliana (thale cress)			
REFERENCE AUTHORS	Arabidopsis thaliana Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi-			
	1 (bases 1 to 741)			
	Castelli,V., Aury,J.M., Jallion,O., Wincker,P., Clepet,C., Menard,M., Cruaud,C., Queietier,F., Scarpelli,C., Schachter,V., Temple,G., Caboche,M., Weissenbach,J. et Salanoubat M.			
TITLE	Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation			
JOURNAL REFERENCE	Unpublished			
AUTHORS	Genoscope.			
JOURNAL	Direct Submission Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)			
COMMENT	The sequences are based on single pass reads. Life Technologies (a division of Invitrogen) members carried out full-length libraries construction ; Temple G. Genosope members carried out sequencing and annotation ; Castelli V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M. URGV INRA : Clepet C., Caboche M.			
	Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences) . 5 prime and 3 prime are assembled with Phrap.			

		http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full1	
		length	
		http://www.genoscope.cns.fr/cgi-bin/ggb/ggb2source=Arabidopsis.	
FEATURES		Location/Qualifiers	
source		1..741	
		/organism="Arabidopsis thaliana"	
		/mol_type="mRNA"	
		/db_xref="taxon:3702"	
		/clone="GSLSIL312B09"	
		/tissue_type="Silique"	
		/ecotype="Col-0"	
		/plasmid="pCWS/PORT_6"	
gene		1..741	
		/gene="At5g42480"	
ORIGIN			
Query Match		26.3%; Score 633; DB 4; Length 741;	
Best Local Similarity		100.0%; Pred. No. 1,1e-167;	
Matches 633; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1774	GCTGGGGCTCAGAGAGTAGCGTTGAGAACTACTGTGAAATGTCGGTGTGATATG	1833
Db	1	GCTGGGGCTCAGAGAGTAGCGTTGAGAACTACTGTGAAATGTCGGTGTGATATG	60
Qy	1834	TTAAAGAGGCAAGTGTGAAGATCTTACGCTGCTGTGGCAATTGACATGATTCAC	1893
Db	61	TTAAAGAGGCAAGTGTGAAGATCTTACGCTGCTGTGGCAATTGACATGATTCAC	120
Qy	1894	TTCAAGCAGAAAGTATTTCTTTAAAGAGCACTCTTTCAAGCAGAAAGTATTTCT	1953
Db	121	TTCAAGCAGAAAGTATTTCTTTAAAGAGCACTCTTTCAAGCAGAAAGTATTTCT	180
Qy	1954	TCTATGGAATCTGATGTGCTACCATAGGGCTAGTCAGAGCTGACGATTCAGAGCA	2013
Db	181	TCTATGGAATCTGATGTGCTACCATAGGGCTAGTCAGAGCTGACGATTCAGAGCA	240
Qy	2014	CCCAAGATGATGCTTAGAGCTGCAAGAGATTTAGTATCCAAATGGCCAGAAATTA	2073
Db	241	CCCAAGATGATGCTTAGAGCTGCAAGAGATTTAGTATCCAAATGGCCAGAAATTA	300
Qy	2074	CTGGCTTTTGGGCGCTGATCAACCGCATAGAAATGTTACCAAGGTTTGGATGG	2133
Db	301	CTGGCTTTTGGGCGCTGATCAACCGCATAGAAATGTTACCAAGGTTTGGATGG	360
Qy	2134	CTGAAGATTTGAGCTGACAGAGCAAGCTGAACTGCGCAGCTTGCGTTTATGATTA	2193
Db	361	CTGAAGATTTGAGCTGACAGAGCAAGCTGAACTGCGCAGCTTGCGTTTATGATTA	420
Qy	2194	ACACTGTTGAACTATCTGTGACAGTGATCAAGTCTCACAAGATGGAACCGGCTCTG	2253
Db	421	ACACTGTTGAACTATCTGTGACAGTGATCAAGTCTCACAAGATGGAACCGGCTCTG	480
Qy	2254	GTCGAAGCAACTCTGAGAGAGTCTGCTGTCTATCTGATTTGGTTTCATCCAGAA	2313
Db	481	GTCGAAGCAACTCTGAGAGAGTCTGCTGTCTATCTGATTTGGTTTCATCCAGAA	540
Qy	2314	GCTACTGATGTCAGAACTTACACAACAAGATGAGATTTTCTGTGTCAAAGTCAG	2373
Db	541	GCTACTGATGTCAGAACTTACACAACAAGATGAGATTTTCTGTGTCAAAGTCAG	600
Qy	2374	AAATCACTGAAGGCTCTGTTCTTGATCATTTAA	2406
Db	601	AAATCACTGAAGGCTCTGTTCTTGATCATTTAA	633
RESULT 4			
AI998415/c		AI998415/c	
LOCUS		561 bp mRNA linear EST 08-SEP-1999	
DEFINITION		701545606 A. thaliana, Columbia Col-0, rosette-2 Arabidopsis	
ACCESSION		AI998415	
VERSION		AI998415.1	
KEYWORDS		GI:5845320	

SOURCE
ORGANISM Arabidopsis thaliana (chale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 561)

REFERENCE
AUTHORS Chen, Y., Momiyama, M., Chan, E., Mooney, M., Carroon, B., Gillingand, D.,
Wang, X., Hillman, J., Guegler, K., Kim, C., Doyle, M., Brzozka, P.,
Gorgone, G., Burns, D., Griffin, J., Mouanoutoua, M., Nguyen, D.,
Tan, R., Rose, M., Warren, B., Ton, B., Kasuriy, K., Borillo, C.,
Cardio, T., Policky, J., Suzuki, G., Argentine, C., Shah, S.,
Nobrega, A., Murry, L., Turner, C., Krikorian, S., Elder, L. and
Hanson, D.
Arabidopsis thaliana Gene Expression Microarray
Unpublished (1999)
Contact: David Smoller, Ph.D.
Genome Systems, Inc., a wholly owned subsidiary of Incyte
Pharmaceuticals, Inc.
463 World Parkway Circle, St. Louis, MO 63134, USA
Tel: 877-577-2733
Fax: 314-427-3324
Email: service@genomesystems.com.

FEATURES
source Location/Qualifiers
1..561
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/ecol_type="Col-0"
/db_xref="taxon:3702"
/clone="701545606"
/issue_type="rosette"
/dev_stage="4 - 7 weeks"
/clone_lib="A. thaliana, Columbia Col-0, rosette-2"
/note="Vector: pSPORT; Site 1: NotI; Site 2: SalI; cDNA
library was derived from untreated rosette tissue from
Arabidopsis thaliana, Columbia Col-0, at 4 - 7 weeks.
Plants were grown in 1:1:1 peat moss/vermiculite/perlite
soil at 22 deg. C +/- 3 deg. C under constant light, and
watered with fertilizer. cDNA synthesis was initiated
using a NotI-oligo(dT) primer. Double-stranded cDNA was
blunted, ligated to SalI adaptors, digested with NotI,
size-selected, and cloned into the NotI and SalI sites of
the pSPORT vector."

ORIGIN
Query Match 20.0%; Score 482; DB 1; Length 561;
Best Local Similarity 99.4%; Pred. No. 7.5e-125;
Matches 482; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
1922 GCTCATCTTTTCAACGCAAGATATGTTCTTCTATGGAATCTGATGCGTACCATAG 1981
Db 561 GCTCATCTTTTCAACGCAAGATATGTTCTTCTATGGAATCTGATGCGTACCATAG 502
1982 GGTGAGTCAGAGTCGAGATTCAGAGCACTCCAGATGAGATGCTAGCACTGACAGA 2041
Db 501 GGTGAGTCAGAGTCGAGATTCAGAGCACTCCAGATGAGATGCTAGCACTGACAGA 442
2042 ATATAGTATCCAGATGCGCAAGATTAAGTCTGCGCTTTTGGGCTGATCACGCAATG 2101
Db 441 ATATAGTATCCAGATGCGCAAGATTAAGTCTGCGCTTTTGGGCTGATCACGCAATG 382
2102 AAATGTTACCGAGAGTTTGGATGGGCGAATGCTGGAAGATTGACTGACAGACAGCTG 2161
Db 381 AAATGTTACCGAGAGTTTGGATGGGCGAATGCTGGAAGATTGACTGACAGACAGCTG 322
2162 AAATGCGGACCTGGGTTGTTATGATTATTAACGTTGAACTATCTGTACAGTG 2221
Db 321 AAATGCGGACCTGGGTTGTTATGATTATTAACGTTGAACTATCTGTACAGTG 262
2222 TGACAGTCTCAGCAGATGGAACCGTGCTCTGTGGAAGCAACTCTGAGAGTCTGCTT 2281
Db 261 TGACAGTCTCAGCAGATGGAACCGTGCTCTGTGGAAGCAACTCTGAGAGTCTGCTT 202
2282 GTCTATCTGATTGGTTGATTCAGAAAAAATGCTACTGATGTGAGAACTTACACACAA 2341

Db 201 GTCTATCTGATTGGTTGATTCAGAAAAAATGCTACTGATGTGAGAACTTACACACAA 142
2342 GATACGAAGTTTGTGTCAGAGTCAGAGGGAATATCATGTAAGGCTGTTTGAT 2401
Db 141 GATACGAAGTTTGTGTCAGAGTCAGAGGGAATATCATGTAAGGCTGTTTGAT 82
2402 CATTA 2406
Db 81 CATTA 77

RESULT 5
B0834167
LOCUS
DEFINITION A1EST0128 Arabidopsis lyrata inflorescence pCMV-PCR library
ACCESSION B0834167
VERSION B0834167.1 GI:28951482
KEYWORDS EST.
SOURCE Arabidopsis lyrata
ORGANISM Arabidopsis lyrata
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 534)
REFERENCE
AUTHORS Barrier, M., Bustamante, C.D., Yu, J. and Purugganan, M.D.
TITLE Selection on rapidly evolving proteins in the Arabidopsis genome
JOURNAL Genetics 163 (2), 723-733 (2003)
PubMed 12618409
COMMENT
Contact: Barrier M
Department of Genetics
North Carolina State University
3513 Gardner Hall, Box 7614, Raleigh, NC 27695, USA
Tel: 919 515 1761
Fax: 919 515 1695
Email: mbarrie@unity.ncsu.edu
Plate: 1 row: H column: 5
Seq primer: T3.
Location/Qualifiers
1..534
/organism="Arabidopsis lyrata"
/mol_type="mRNA"
/cultiivar="Kathumaki"
/db_xref="taxon:59689"
/clone="PIWB1-D03"
/issue_type="inflorescence"
/clone_lib="Arabidopsis lyrata inflorescence pCMV-PCR
library"
/note="Vector: pCMV-PCR (Stratagene); Created using PCR
library Construction kit (Stratagene)"

ORIGIN
Query Match 19.0%; Score 458; DB 5; Length 534;
Best Local Similarity 94.1%; Pred. No. 4.8e-118;
Matches 476; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
1901 AGAAGTATTTTCTTAAAGAGCTCATCTTTTCAAGCAAGATATGTTCTTCTAATG 1960
Db 2 AGAAGTATTTTCTTAAAGAGCTCATCTTTTCAAGCAAGATATGTTCTTCTAATG 61
1961 AATCTGATGTCGCTACCAATAGGTCAGTCAAGTCGATTCAGAAAGCACTTCCAGAA 2020
Db 62 AATCTGATGTCGCTACCAATAGGTCAGTCAAGTCGATTCAGAAAGCACTTCCAGAA 121
2021 TGATGCTAGAGCTGCGAGAAATTAATATTCAGAAAGTGGCGAAAGATTAAGTCTTGCTT 2080
Db 122 TGATGCTAGAGCTGCGAGAAATTAATATTCAGAAAGTGGCGAAAGATTAAGTCTTGCTT 181
2081 TTGGGCTGATTCACCGATGAATGTTTACAGAGTTTGGATGGCGAATGCTGAGA 2140
Db 182 TTGGGCTGATTCACCGATGAATGTTTACAGAGTTTGGATGGCGAATGCTGAGA 241

Oy	2141	TTTGGACGTACACGACGACTGAAATCTGGCCAGCTGGGTGGTTATGATATACACTGT	2200
Db	242	TTTGGACGTACACGACGCTGTGGAAATCTGGCCAGCTGGGTGGTTATGATATACACTGT	301
Oy	2201	TGAACCTATCTGTTACAGTGTGACAGTCTGACAGATGGAACCCGCTCTGGTGGAG	2260
Db	302	TGAACCTATCTGTTACAGTGTGACAGTCTGACAGATGGAACCCGCTCTGGTGGAG	361
Oy	2261	CAACTCTGGAGAGTCTGCTGTTCTATCTGATTTTGGTTATCTCAGAAAACATGCTACTG	2320
Db	362	CAACTCTGGAGAGTCTGCTGTTCTATCTGATTTTGGTTATCTCAGAAAACATGCTAAAG	421
Oy	2321	ATGTCAGAACCTACACAACAAGATACGAAGTTTCTGTCACAGTCAAGGTGGAAAAATCA	2380
Db	422	ATGTCAGAACCTACACAACAAGATACGAAGTCTGTCACAGTCAAGGTGGAAAAATCA	481
Oy	2381	CTGAAGGCTCTGTTCTTGATCATCATTA	2406
Db	482	CTGAAGGCTCTGTTCTTGATCATCATTA	507
RESULT 6			
CL965374		2307 bp	DNA
LOCUS			linear
DEFINITION	OsIFPC012338	Oryza sativa Express Library Oryza sativa (indica cultivar-group) genomic, genomic survey sequence.	
ACCESSION	CL965374		
VERSION	CL965374.1	GI:52385433	
KEYWORDS	GSS.		
SOURCE	Oryza sativa (indica cultivar-group)		
ORGANISM	Oryza sativa (indica cultivar-group)		
REFERENCE			
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.		
	1 (bases 1 to 2307)		
	Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M., Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L., Wong, G. K. S., Deng, X. W. and Wang, J.		
	An analysis of transcriptional regulation of the rice genome and its comparison to Arabidopsis		
	Unpublished (2004)		
JOURNAL	Contact: Chen Chen		
COMMENT	Department of Bioinformatic		
	Beijing Institute of Genomics		
	Chinese Academy of Sciences, Beijing 101300, China		
	Tel: 86-10-80481559		
	Fax: 86-10-80488676		
	Email: chenchen@genomics.org.cn		
	Rice genomic sequence.		
	Class: exon-trapped.		
FEATURES			
source	1..2307		
	location/Qualifiers		
	/organism="Oryza sativa (indica cultivar-group)"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:39946"		
	/clone_idb="Oryza sativa Express Library"		
	/note="Oryza sativa exon trapped genomic sequences "		
ORIGIN			
Query Match	18.6%;	Score 447;	DB 10;
Best Local Similarity	53.7%;	Pred. No. 1e-114;	
Matches 1256;	Conservative	0;	Mismatches 880;
		Indels 201;	Gaps 8;
Oy	108	CTGCTCGCCGACGAATAGGCGACCGCTCTCTCCGACTTGAATTTACACCTCCGATTC	167
Db	132	CTGGGCGGAACGCTCTTCCGCACTTCCACACTCTCTCCACCGCGCGCTCTCGACCC	191
Oy	168	CTCTCTCTCTCTTCTGGCACCGCACCAACCGCACTCTCTCTCTTCTGCACCATC	227
Db	192	GCGCTCCCGGCTT	251
Oy	228	TATTGATCTCCGACGACGCTCCCATCTCCATTTGATTTTCAACAGGAT-----	280

Db	252	CCCCGACGCCGCCGAAACGCTCCCTCCGCTCCAACTGCAATTTCTTCAAAAGTGAGGCCGCCG	311
Oy	281	-----TAGAGCTCAAAACATTTCTTAAACCGATGGAAATCAGAAAGAC	323
Db	312	TGAATCTGCTCCGCTTCTAGAGGCGAGAGCCACATTTCTTGGCGATGGCAATCAGAGGAC	371
Oy	324	ATTGGAAGCTAAGGTTTGGAAAACCGCGCAATTCGGTTTACGCCAGCAGCGCTTAAATCAG	383
Db	372	GTTGAGGCAAGGAAATGACCAAGCCACCGCAGATATGGCTACAGCAACGAAATGCTCTGTTGG	431
Oy	384	CCGAGACAGATTTCTTCAAGCTGCTTGGAAACTCTGCTTAATCTCGGTCTAGAAAGA	443
Db	432	TGCTGCAAAATGCTGCAGATTGGCCCATGACACTCATGAAACCAAGACTCCCGCATCTCA	491
Oy	444	GTAACAATGAAGTCTTCTTGATGATGAAGAAGCTACAGTCATCACTGATGTTCTCTGGGA	503
Db	492	GTAATGATGTCGCTTCTTGAGAACCGGAAAGAACTTCAACATGATATTTGCTTGGGA	551
Oy	504	TAAAGTTCTGAGGCTCTGTGTGATATGCAAGAAAGTGTGAGACTGAGATAGTTCTTTCG	563
Db	552	CAAGGAGGCTGG-----GAGGCACTTGCTGTGCTTGT	584
Oy	584	GGTGTGTGAGGCTCTGCTTAAGGAGAGGTTCCTTAAGCTTTAAAGCAAGATGTGTTT	623
Db	585	AACGTGAAMACAGTTGCTCTTGATATCGGCCACCCAAAGGCTTCAAGCAGACGTGTGCT	644
Oy	624	AGTTATGGCGCTTGCGTTTTCGATATGTCGAGGGAATGCTATGGAATTGATTCACCTGA	683
Db	645	AGCATGCTCTTGCTTATGTGATCTATCAAGGAAATCTATGGCAGAGACCTTCCAGA	704
Oy	684	TTTTATTACTGTTATGAGTTTGTGAGGAAGCTTGAAGCTTTTACAGAGGAAGAGC	743
Db	705	TGTAATTTGCTGTGCGAGGTGCTCGAAGGGCTCTCAAGCTCTTGCGAGAAATGAGC	764
Oy	744	AAATGACCTTGACCCGAAATTTACGTGACAAATTTGATGAGACTTTGGAAAGATCACTCC	803
Db	765	AAAGATCTCGCAACTGATCTGCTTTACAGATTAAGAACTCTCGAGGAAATTAACCC	824
Oy	804	GCGTTATGCTCTGAGGCTAATCGGCTTACGCTTGATGATTAACGCTGCGAAAGACT	863
Db	825	TGCGTGTATTTGAGGCTTCTCTCCCTTCCATTGACACAGAGCATATAAGAAAGCCCA	884
Oy	864	AAATGTTTAAACCGATGTGCGGAATATTTGTGCTGTGAGGAGGTGAGCATCAGC	923
Db	885	AGAAAGGCTTCAAGGTGGAGAAACATTTGTGTGAGGTTGGCAGAGAGGTATTTGCTAC	944
Oy	924	TCTTGTGGGGGTTTGAACCGGTGGAAGTTTATGATGAGCGGTTTTTACGATGACGC	983
Db	945	CGTTGAGAGGAAATTTCTCGTGAAGCCTTATGAAGAGGCTTTTGTGAGGATGACATC	1004
Oy	984	TGCTGACAGGTTGATCTTTTGTATGCTAACCCCAAGCAATATTCACACAGTCATTTGA	1043
Db	1005	AATTGAACAGATGATTTCTTTCAAAAACCGAATAGCATTCCTCTGATATGTTTGA	1064
Oy	1044	AGTTTAAAGTTGCACTTGCTCTTGTGAGCTCAAGCTTTTAAATTTGTAAGAACCAACT	1103
Db	1065	AATTTAAATGACACTTGACATGTGCGTCAAGCAATTAATGATTAAGGCCACATTT	1124
Oy	1104	TTTACAGAGTGTGATTAAGCAATTTCCAGCACTTCAGACGCTAAGSTAAATGCTATGGA	1163
Db	1125	CATCATATAGCGGATATCTTTTGTGAACAATCCAAAGTTTCA-----	1169
Oy	1164	GATTCCTGCGATGTTGATGATACAGGAATTAATGGAGATGACTTCGGTCAAGAAAG	1223
Db	1170	-----CATAGTTCTCATTTATGCTTATGATATATGAGATGACCTTGCAATTTGAAG	1220
Oy	1224	GGGACTGTGCACTGCTTAAAGCAAAAGTTGATGATGCGTATATGTGTTGGCTTTAGA	1283
Db	1221	GGCAATCTGCTCAATGTCTAGTCGGAAGATGTTAGCAATGCAAGATGTGGCTTGAATTGA	1280
Oy	1284	CAGTGAAGATTCACAAATATGGAATTCAGCTATTTGTGAGATTGTTTGGAGAAATTC--	1340
Db	1281	TAAAGATCTTCAACCATACAGAGCCCAAAATTTCTAGATTTATGTGACCAACTCTAG	1340

1341 AAATGATGACATGATGATCTCCCTGACTATGCAAAATTGTTGAAAACCTGGTGGC 1400
1341 CATGAGTAAGATGATCTCTCCAGGGCTGTGCAAGCTTTTGAGACTTGGCTTAT 1400
1401 AGGGGTCTCTTCTAGGTTGAGAGACCAAAAGATTAATAAATTAACTCGGGACTA 1460
1401 CTTTGGGTTTCTTAGAGACAGATATCTGGGGCATGTGAGTTGACACTTGAAGATTA 1460
1461 CTATGATGATCTTATGTTTGAATTACTTGAAAAGATGAGGATGTTCAAGGTTCTCC 1520
1461 CTAGCATATTCAGAAAGTTTAAAGCTTAAAGATGAGAGGTTGGTGTCTTCTCA 1520
1521 TTTAGCTGCTGCTGCACTATGAGAGGATTTGAGCCGACATGTAAGCTAGTCTAT 1580
1521 TTTGGCTGCTGCTGCTATTTGAGAAACCTTGCTCAGCTACAGCTGACCTGGTA- 1579
1581 GCAGGCACTGAGAAAGTTTCTTCCCTCCGCTATACAGATAGAACTCGCTGAACCA 1640
1580 -----CTGGAATCMAA 1592
1641 GGATGTGCAAGACAGTGTATGATGATCTGTGTGAACAATGAGCCGTGATGG 1700
1593 TGCTATTCAAGCG----- 1605
1701 TGAGCTGTGTCTTTATTGCAAGCTGTAAAGCCCTGTGAAAACCTTGAACATAATGA 1760
1606 -----TTCAACAAGGTTTTCATGATGATAGACAGTTAGACAGGTGACG 1649
1761 TTTATGCAATTCAGAGCTGGGGTCTCAGAGATGAGCTGTATGATAAACTGTGAATGTC 1820
1650 CATGAAAATCTAAAGATGAGCCCTGGGGGATATCTTGAATAATTTTGACAGAAAATGC 1709
1821 CGTTGCTGATATGTTAAAGAGGAGCAAGTGAAGATCTTACTGCTGTGTGTGCAATTGG 1880
1710 ACTGCTATATTTGAGAAATGCGCCTTGAAGATTAATCTGTGCTGCGCACTGTTGC 1769
1881 ACTGATTTCACTGTTCAACGCAAGATATTTTCTTAAAGACAGCTATCTTTCAACGCA 1940
1770 ACTGTGACAGTAATTTGGGGCAATAATTTGCTGTGAAGAGGCCCTTTGCTGATTAAG 1829
1941 GGAATGATTTCTTCTATGAAATCTGAT-----GTGCTACCATAGGGTCACTCAG 1991
1830 GAGTGAAGATGATCTGTGGAGTGTCTAATAGTGTCACTTACTGATGATCTGCACT 1889
1992 AGCTGACATTTCAAGAGCACTTCCAGAAATGAGATCTGAGACTGAGAGAAATATGATATC 2051
1890 AGATGAGATTCAGTACATATTTCTAGAAATGATGCAAGCTGGCAGAGAAATTTGTCG 1949
2052 CAAGTGACAGAAATTAAGTCTTGGCTTTTGGGCTGTATGACCGCATGAAATGTTACC 2111
1950 CAAGTGACAGAAATTAAGTCTTGGGCTTGGGACCAAGATCTGGTTCATCTTCA 2009
2112 AGAGTTTGAATGGGCGAATGCTGAAGATTTGCACTGACAGAGAGCTGAAATGCGCA 2171
2010 AGAGTTTCTTATGCAACATGCTTAAAGTGTGACCTGACGAGAGGAGAGATTTGACCG 2069
2122 GCTTGGTGTGTTTGAATGATTAACCTGTTGAACTATCTGTGACATGAGAGTCTC 2231
2070 TCAATGGTGTCTTGGAGTATACATCCGATGTGACATGATGATCACTATCTC 2129
2232 AGCAGATGAAACCCGTGCTGTGTGGAAGCACTCTGAGAGAGTGTGCTTCTATCTGA 2291
2130 CCTAGATGGTGAAGAGGACGCTGTGAGGCTACGATGATGAGGAGGCAACTTACTGA 2189
2292 TTTGGTTCATCCAGAAACATGCTACTGATGTGCAAACTTACACAGACAGATGCGAAGT 2351
2190 TGTTAAGAGCCAGAAACATGATTCATATGACAAATAACTACCCGATGATGAT 2249
2352 TTTCTGCTCAGAG---TCAGGGTGAATAATCAGAGGCTGTGTTCTTGATATA 2405
2250 GAGCTTCTCAAGCTTGAAGAGGAGGAGATTAAGAGAGAGAGTCTCAAGTGTGA 2306

RESULT 7
CO071968 813 bp mRNA linear EST 15-JUN-2004
LOCUS GR_Ea30N03.r GR_Ea Gossypium raimondii cDNA clone GR_Ea30N03 3',
DEFINITION mRNA sequence.
ACCESSION CO071968
VERSION CO071968
KEYWORDS GI:48741449
SOURCE EST.
ORGANISM Gossypium raimondii
Gossypium raimondii
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
1 (bases 1 to 813)
REFERENCE
AUTHORS Kam,H., Yu,Y., Kudrna,D., Hatfield,J., Stum,D., Mueller,C.,
Udall,J.A., Rapp,R.A., Mendel,J.F., Rao,K., Soderlund,C. and
Wing,R.A.
TITLE Global assembly of Cotton ESTs
JOURNAL Unpublished (2004)
COMMENT Contact: Rod A. Wing
Arizona Genomics Institute
The University of Arizona
Forbes Building Room 503, Tucson, AZ, 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: rwing@genome.arizona.edu
Plate: 30 row: N column: 03.
FEATURES
source
Location/Qualifiers
1..813
/organism="Gossypium raimondii"
/mol_type="mRNA"
/db_xref="taxon:29730"
/clone="GR_Ea30N03"
/tissue_type="whole seedlings"
/dev_stage="first true leaves"
/lab_host="DH10B"
/clone_lib="GR_Ea"
/note="Vector: PCMV.SPORT-6.1; Site 1: NotI; Site 2:
EcoRV; Library made by Invitrogen with RNA supplied by
Wendle lab. Directional cloned into NotI-EV. Colonies
plated/picked by AGT. More glycerol clones held in -80."

ORIGIN
Query Match 18.1%; Score 436; DB 7; Length 813;
Best Local Similarity 72.3%; Pred. No. 9,7e-112;
Matches 584; Conservative 0; Mismatches 215; Indels 9; Gaps 1;
107 TCTGCTCCGCGAGAAATGGGCGGACCGCTTCTTCCGACTTCAATTTCACTCG--- 163
5 TCTGCTCCGCGAGAAATGGGCGGACCGCTTCTTCCGACTTCAATTTCTCCCTGCCC 64
164 -----ATTGCTCTCCTCCTCCTGCGCAACCGGCAACCAACCGGCACTGTGTCCTC 217
65 CCGATATTCGTCCTCTCTTCTTCTTCTCAACCGGCACTTCTTCTCCCTTTACCTTC 124
218 TGCACCATCTATTTGATCGTCCGGAAGCCGACGTCCTCCCATCCCATTTGATTTTACAGG 277
125 CCTTTCGCGCCCTTCTCCTCCGGAAGCCGACGTTCCATTTCCCTTTGATTTTCAAGG 184
278 TATTAGAGCTCAACACATTTCTTAAACGATGATCAGAAAGCACTTTGAAAGCTAAGG 337
185 TTTTGAAGCCGAGACTCATTTCTTATGATGATGAAATCAAGAAAGCCTATGAAGCAAGG 244
338 TTTTGAAGCCGCGCAATTTGGTTTGAAGCAAGCCTTAAATCAAGCCGAGACAGATTTC 397
245 TTTTGAAGCCGCGCAATTTGGTTTGAAGCAAGCCTTAAATCAAGCCGAGACAGATTTC 304
398 TTTCAAGCTGCTTGAAGAACTCTGTCTAATCTCGGCTTGAAGAGAGTACATGAAAGTCT 457
305 TTTCAAGCTGCTTGAAGAAACCTTATTAACCTGCTTGAAGAGAAATTAACCAAGAGTCT 364
458 TTTTGAATGATGAAGAGCTACAGTCACTGATGTTCTTGGGATTAAGGTTCTTGGG 517

Db 365 TTGTGAGAGTGAAGCGTGAACATATCATCATCAAGTCCCTGGAGACAAGTTCTTGAG 424
Oy 518 CTCTCTGTATTTGCAAGAGGTTGGTGAAGTGAATGTTCTTGGGTTGGTGAAGCTC 577
Db 425 CATCTGCTGTGTGCAAGAGTGGGAGACGTAAGTGTGCTTCAATTTGGGAGAGTT 484
Oy 578 TGCTTAAGAGAGTGTGCTTAAGTCTTTAAGCAAGATGTGTTTAACTTAATGCGCTTG 637
Db 485 TGCTGAGAGAGAGGCTGCCCAAGCATTTTAAACAGAGTGTGCTTGGCAATGCGCTTG 544
Oy 638 CGTTCTGTAGTCTGAGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 697
Db 545 CTATGAGACTTGTGAGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 604
Oy 698 ATGAGTTGTTGAGAGACCTTTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 757
Db 605 GTGAGTCTGTAAGAGGCTTTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 664
Oy 758 CGAGTTTACGTGACAAATTAATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 817
Db 665 CGAGTTTACGTGACAAATTAATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 724
Oy 818 AGCTACTGCTTACCGCTTGTGATGAATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 877
Db 725 AACTTCTGACCTTGTGCTTGTGATGAATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 784
Oy 878 GGTGCGGAGATTTTGTGCTGTTGG 905
Db 785 GTGTAGCGACATTAATTAAGGCTGTGG 812

RESULT 8
CD573714 897 bp mRNA linear EST 12-JUN-2003
LOCUS UCRPT01_01_F12_T3 Poncirus trifoliata CTV-challenged cDNA library -
DEFINITION UCRPT01_01_F12_T3 Poncirus trifoliata cDNA clone UCRPT01_01_F12, mRNA sequence.
ACCESSION CD573714
VERSION CD573714.1 GI:3169616
KEYWORDS EST.
SOURCE Poncirus trifoliata
ORGANISM Poncirus trifoliata
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Sapindales; Rutaceae; Poncirus.
1 (bases 1 to 897)
REFERENCE 1
AUTHORS Roose, M.L., Ye, X., Federico, C.F., Close, T.J., Fenton, R.D.,
Wanamaker, S., Choi, Y. and Kingan, T.
TITLE Development of EST Resources and New Genetic Markers for California
JOURNAL Citrus - Poncirus trifoliata CTV-challenged phloem - UCR
COMMENT Unpublished (2003)
Contact: Mikeal Roose
Department of Botany & Plant Sciences, University of California
Riverside, CA, 92521-0124, USA
Tel: 909/874137
Fax: 909/874437
Email: mikeal.roose@ucr.edu
Seq primer: T3.
FEATURES
Source
1..897
Location/Qualifiers
/organism="Poncirus trifoliata"
/mol_type="mRNA"
/cultivar="Pomeroy OP"
/db_xref="taxon:37690"
/clone="UCRPT01_01_F12"
/tissue_type="Phloem"
/dev_stage="10 - 30 cm shoots"
/lab_host="E. coli TJC121"
/clone_lib="Poncirus trifoliata CTV-challenged cDNA
library - UCR"
/note="Vector: Lambda Uni-ZAP XR, excised phagemid,
Site 1: EcoRI, Site 2: XhoI. Plants were grown in the
greenhouse at University of California, Riverside. The

ORIGIN

scion was a open-pollinated (very probably selfed)
seedling of Poncirus trifoliata cv Pomeroy that was
selected as homozygous for the Ctv resistance gene. The
rootstock was sweet orange infected with citrus tristeza
virus (CTV) isolate 7514 over 1 year before sampling (CTV
infects sweet orange, but not genotypes carrying the Ctv
resistance gene. Shoots 10-30 cm long were harvested in
October 2000, and the green phloem (bark) was removed and
frozen quickly in dry ice. Total RNA was extracted using
Trizol reagent (Gibco). Poly(A) RNA was purified, a cDNA
library was made, and 0.5 million primary lambda cDNA
clones were in vivo excised to give a population of
phagescript SK(-) phagemids. All steps to this point were
performed in the ML Roose lab at the University of
California, Riverside by X. Ye. Phagemids were plated,
plasmid DNA purified, cDNA clones archived, and DNA
sequences determined bi-directionally using an ABI730 at
the University of California Riverside Genomics Institute,
Core Instrumentation Facility, (Choi, Kingan).
Chromatogram files were transmitted to UC Riverside (by
Choi), then processed at UC Riverside (by Wanamaker) using
the Harvest pipeline (<http://harvest.ucr.edu>) to remove
vector and cloning oligo sequences and various
contaminants, and to trim to a high quality region.
Sequences that retained a phred 17 region of at least 100
bases were deposited to Genbank."

Query Match 17.8%; Score 429.4; DB 6; Length 897;
Best Local Similarity 70.4%; Pred. No. 7.5e-110;
Matches 574; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

Oy 742 GCAAGTAGCCTTGACCGAGATTTTACGCAAAATTAATGAAGTGAAGTGAAGTGAAGTGAAGT 801
Db 80 GCCAGTAGCTTCGCTCCAGATTTTGACGACACAGTGAAGTGAAGTGAAGTGAAGTGAAGT 139
Oy 802 CCGGTTATGCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 861
Db 140 CCACGCTGTGTTGGAACCTTTAGGCTTACCCCTTGTGATGATGATGATGATGATGATGATG 199
Oy 862 CTAAATGTTTAACCGGTGCGGAATTTTGTGCTGTGATGATGATGATGATGATGATGATG 921
Db 200 GAAAGAGACTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 259
Oy 922 GCTCTTGTGGGGGTTTGAACCGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 981
Db 260 GCATTTGTGGGGATTTTACGCTGAATCTTCAATGAATGAGGCTTTTCAATGATGATGATG 319
Oy 982 GCTGCTGACGAGTGAATCTTTTGTGATGATGATGATGATGATGATGATGATGATGATG 1041
Db 320 TCAGCTGAGCAGGTTAACTTTTGTGCGCACCAATTAATGATGATGATGATGATGATGATG 379
Oy 1042 GAACTTACGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1101
Db 380 GAACTTACGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 439
Oy 1102 CTTTACGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1161
Db 440 CTGATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 499
Oy 1162 GAGATTCCTGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1221
Db 500 CGGAGCTTGGCTTATCTATATTCCTTGGAAAAACATGATGATGATGATGATGATGATGATG 559
Oy 1222 AGGGAGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1281
Db 560 AGGGAGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 619
Oy 1282 GACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1341
Db 620 GATG 679
Oy 1342 AATG 1401

Db 680 AAGGAAGCTAATGACATGATCTTCTGATCTCTGTAAGCTGTAGACATGCGTGACA 739
Oy 1402 GGGGTGCTCTTCCCTAGCTTGCAGACACCAAAAGTAAATAATTAACTGGGAGACAC 1461
Db 740 GAGGTGTTTCCCTAGATTCAGGGAGACCAAGTGAATACGGTTCAACTGGTGATTAAC 799
Oy 1462 TATGATGATCTTATGATTTTGGAGTACTTGGAAAAGTGAAGTACTCAGGGTTCTCT 1521
Db 800 TATGATGATCTTACTGCTCTTCTGATTTTGAAGAAACATCGAAGGAAATGATGTTCCGC 859
Oy 1522 TTAGCTGCTGTGCAACTATGGAAGATTGAGC 1556
Db 860 TTAGCTGACAGAGCTGCATAGTATTAGGCTTGAGC 894

RESULT 9
B2437564/c B2437564 785 bp DNA linear GSS 13-DEC-2002
LOCUS BONRN72TR.B0.1.6.2 KB tot Brassica oleracea genomic clone BONRN72,
DEFINITION genomic survey sequence.
ACCESSION B2437564 GI:26691135
VERSION B2437564.1 GI:26691135
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 785)
Aylee,M., Haas,B.-J., Kumar,N., Wu,H., Xiao,Y., Van Aken,S.,
Utecherbck,T.R., Wortman,J.R., White,O.R. and Town,C.D.
Whole genome shotgun sequencing of Brassica oleracea and its
application to gene discovery and annotation in Arabidopsis
Genome Res. 15 (4), 487-495 (2005)

JOURNAL
PUBMED
COMMENT
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtow@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TR
Classes: sheared ends.
Location/Qualifiers
1..785
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="T01000DH3"
/db_xref="taxon:3712"
/clone="BONRN72"
/clone_1lb="B0.1.6.2 KB tot"
/note="Vector: PHOS1; Site_1: BactXI; 1.6-2 kb sheared
total DNA inserted into PHOS1 using BactXI linkers"

ORIGIN
Query Match 15.0%; Score 360.8; DB 9; Length 785;
Best Local Similarity 86.4%; Pred.No.2.1e-90;
Matches 425; Conservative 0; Mismatches 57; Indels 10; Gaps 2;

Oy 505 AAGGTCTCTGGGAGCTCTGTTATTCAGAGAAGGTGTGAGACTGAGATAGTTCTTGG 564
Db 693 AAGGTCTCTGGGAGCTCTGTTATTCAGAGAAGCTGTGAGACTGAGATAGTTCTTGGT 634
Oy 565 GTTGTGAGGCTCTGCTTAAGAGAGTTGCCCTAAGTCTTAAGCAAGATGTGGTTTAA 624
Db 633 GTAGAGAGAGCTTCTTAAGGAGAGTTGCCCTAAGTCTTCAAGCAAGATGTGGTTTGG 574
Oy 635 GTTATGAGGCTTGTCTGATGTCTGAGAGATGTATGAGCATTTGGATTCACCTAT 684
Db 573 GTTATGAGGCTTGTCTGATGTCTGAGAGATGTATGAGCATTTGGATTCACCTAT 514

Oy 685 TTATTAAGCTGTTATGAGTTGTTGAGAGAGCTTTGAAGCTTTTACAGAGAGAGACA 744
Db 513 TTATTAAGCTGTTATGAGTTGTTGAGAGAGCTTTGAAGCTTTTACAGAGAGAGAGACC 454
Oy 745 AGTACCTTGACCGGATTTTACGTGACCAATATGATGAGACTTTGGAAGATCATCTCG 804
Db 453 AGCAGCTTGACCTGATTTTACCGGCCAGATGATGAGACTTTGGAAGATCATCTCG 394
Oy 805 C-GTTATGCTTGGAGTACTGCGCTTACCGCTTGAGATGATTAAGCTCGAAAGACT 863
Db 393 CGTTATGATATGAGCTTCTGTTTACCTCTGGGGAGATGAT-----AAACACA 343
Oy 864 AAATGTTTAAAGCGGTGCGAATATTTGTTGCTGTTGAGAGAGTGGAGCATCAGC 923
Db 342 AGATGTTTAAAGCGGTGATGAGCAATATCTGTGCTGTTGAGAGAGCGGAGCATCAGC 283
Oy 924 TCTTGTGGGAGTTTGAACCGGTGAGAAATTATGATGAGCGCTTTTACGAATGACAGC 983
Db 282 TGTTGTGGTGTCTTTCACGTGAGAAATTTATGATGAGCGCTTTTACGAATGACAGC 223
Oy 984 TCGTACGACAGT 995
Db 222 ATCTGACGAAGT 211

RESULT 10
CX195205 624 bp mRNA linear EST 28-DEC-2004
LOCUS 13-EP19365-021-009-J03-TR ADIS-MP12 021 Brassica napus cDNA clone
DEFINITION MP12p102J039Q, mRNA sequence.
ACCESSION CX195205
VERSION CX195205.1 GI:56842629
KEYWORDS EST.
SOURCE Brassica napus (rape)
ORGANISM Brassica napus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 624)
Jakoby,M., Lehmann,D. and Weisshaar,B.
direct submission to Genbank (ADIS-MP12 021)
Unpublished (2004)
Contact: Bernd Weisshaar
Chair of Genome Research
Bielefeld University, Institute for Genome Research
Universitaetsstrasse 25, D-33594 Bielefeld, Germany
Tel: +49-521-106 6873
Fax: +49-521-106 6423
Email: bernd.weisshaar@uni-bielefeld.de
Insert Length: 624 Std Error: 0.00
Plate: 9 row: J column: 3
Seq primer: T7R CTATATGACTCATATGAGGA.
Location/Qualifiers
1..624
/organism="Brassica napus"
/mol_type="mRNA"
/cultiVar="Express 617"
/db_xref="GABI:1111775"
/db_xref="taxon:3708"
/clone="MP12p102J039Q"
/tissue_type="flowers and flower buds"
/dev_stage="young flowers and flower buds"
/clone_1lb="ADIS-MP12 021"
/note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; cDNA
library from Brassica napus, strain Express 617; RNA was
from young flowers and flower buds of two greenhouse
plants; library was made at the Max-Planck-Institute for
Plant Breeding Research, Cologne, Germany; cloning sites
SalI-NotI, primer sites and orientation:
T7-salI-CCACGCGTCCG-5err-cDNA-polyA-CC-NotI-Spe; made by
Marc Jakoby 09/2000; PI: Bernd Weisshaar. Sequence
submission managed by RZPD/GABI-Primary database:

ORIGIN http://gabi.rzpd.de"

Query Match 14.9%; Score 359.2; DB 8; Length 624;
Best Local Similarity 76.9%; Pred. No. 5.6e-90;
Matches 518; Conservative 0; Mismatches 93; Indels 63; Gaps 4;

```
OY 1303 AGAATCCAGCTATTGTGAGAGTTTGTGAGAAATTCGAATCGTGATGACATGATGAT 1362
DB 2 AGGAATCCAGCTATTGTGACCTTGTCTTGAGAAATTCGAATCGTGATGACATGACGAC 61
OY 1363 CTCCTCGGCTCTGCAAAATTTAGAGACCTGGTGGCGGGGTCTCTTCTAGAGTTT 1422
DB 62 CTCCTCGGCTCTGCAAAATTTAGAGACCTGGTGGCGGGGTCTCTTCTAGAGTTT 121
OY 1423 AGAGACACCAAGATTTAAATTTAACTCGGGGACCTACATGATGATGATGATGATGAT 1482
DB 122 AAGATACCAAGATTTAAATTTAACTCGGGGATTTAAGATGATGATGATGATGATGAT 181
OY 1483 AGTTACTTGAAAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1542
DB 182 AGTTACTTGAAAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 241
OY 1543 GGAAGATTTGAGCCGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1602
DB 242 GGAAGATTTGAGCCGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 301
OY 1603 CCTTCCCGCTATACAGATGAAATCGGCTGAACCCAGATGTCAGAGACAGTGT 1662
DB 302 CCTTCCCGCTATACAGATGAAATCGGCTGAACCCAGATGTCAGAGACAGTGT 337
OY 1663 AGTGTAGATCTGTGTGTAACATGAGCCGTA---TGAGAGCTGTGTCTTAT 1719
DB 338 ACTGTAGATCTGTGTGTAACATGAGCCGTA---TGAGAGCTGTGTCTTAT 397
OY 1720 GAGAGAGCTGTGTGTAACATGAGCCGTA---TGAGAGCTGTGTCTTAT 1779
DB 398 GAGAGAGCTGTGTGTAACATGAGCCGTA---TGAGAGCTGTGTCTTAT 431
OY 1780 GTCTCAGAGAGTGTGTGTAACATGAGCCGTA---TGAGAGCTGTGTCTTAT 1839
DB 432 ---GAAAGTGTGTGTGTAACATGAGCCGTA---TGAGAGCTGTGTCTTAT 487
OY 1840 GAGGCAAGTGTGTGTAACATGAGCCGTA---TGAGAGCTGTGTCTTAT 1899
DB 488 GAGGCAAGTGTGTGTAACATGAGCCGTA---TGAGAGCTGTGTCTTAT 547
OY 1900 CAGAGATTTTCTTAAAA-----GCACTCATCTTTTCAAGCAAGATGATGAT 1953
DB 548 CTGAGATTTTCTTAAAA-----GCACTCATCTTTTCAAGCAAGATGATGATGAT 607
OY 1954 TCTATGATGATGAT 1967
DB 608 TCTATGATGATGAT 621
```

RESULT 11
LOCUS CO075595
DEFINITION GR_Ea36G17.r GR_Ea Goeypium raimondii cDNA clone GR_Ea36G17.3', mRNA sequence.
ACCESSION CO075595
VERSION CO075595.1 GI:48745076
KEYWORDS EST.
SOURCE Goeypium raimondii
ORGANISM Goeypium raimondii
Bukeriyota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Goeypium.
REFERENCE 1 (bases 1 to 769)
Kim,H., Yu,Y., Kudrna,D., Hatfield,J., Stum,D., Mueller,C., Udall,J.A., Rapp,R.A., Wendel,J.F., Rao,K., Soderlund,C. and Wing,R.A.

TITLE Global assembly of Cotton ESTs
JOURNAL Unpublished (2004)
COMMENT Contact: Rod A. Wing
Arizona Genomics Institute
The University of Arizona
Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: rwing@genome.arizona.edu
Plate: 36 row: G column: 17.

FEATURES
source location/Qualifiers
1..769
/organism="Goeypium raimondii"
/mol_type="mRNA"
/db_xref="taxon:29730"
/clone="GR_Ea36G17"
/tissue_type="whole seedlings"
/dev_stage="first true leaves"
/lab_host="DH10B"
/clone_id="GR_Ea"
/note="Vector: PCMV-SPORT-6.1; Site 1: NotI; Site 2: EcoRV; Library made by Invitrogen with RNA supplied by Wendle lab. Directional cloned into NotI-EV. Colonies plated/picked by AGT. More glycerol clones held in -80."

ORIGIN

Query Match 14.3%; Score 344.2; DB 7; Length 769;
Best Local Similarity 68.9%; Pred. No. 1.1e-85;
Matches 491; Conservative 0; Mismatches 213; Indels 9; Gaps 1;

```
OY 1 ATGAAAGCTGTGAGTCAAGTGGGCAATGATGATGATGATGATGATGATGATGATGAT 60
DB 54 ATGAAATCTTGAAGCATATGATGATGATGATGATGATGATGATGATGATGATGAT 113
OY 61 CCGGGAAGCAAAAGCTCCGAGGTAAGCCAAACACTTACACTATGCTGCTCCGAC 120
DB 114 CACCTCTTAAACCTCCAGACTCCACCGCTTCACTACCGTGTGCTGCTGCGCAGT 173
OY 121 AAATGGGCGAAGCTCTCTCCGACTTCAATTT-----CACCTCGATTCCTCC 171
DB 174 AAATGGGCGAAGCTCTCTCCGACTTCAATTT-----CACCTCGATTCCTCC 233
OY 172 TCTCTCTCTTCCGCAACGCGCAACCAACGCGCACTCTGCTCTCTGCAACATCTAT 231
DB 234 TCTCTCTCTCTCTCTCAACGCGCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 293
OY 232 GATGCTCCGAAAGCGCAAGTCCCACTCCCACTGATTTCTACAGATTTAGAGCTCAA 291
DB 294 TCTCTCTCCGAAAGCGCAAGTCCCACTCCCACTGATTTCTACAGATTTAGAGCTCAA 353
OY 292 AACATTTCTTAACCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 351
DB 354 ACTCATTTCTTAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 413
OY 352 CAATTCGTTTCAAGCAAGAGCTTTTATTCAGCCGAGACAGATTTCTTCAAGCTGCTGC 411
DB 414 CAATTCGTTTCAAGCAAGAGCTTTTATTCAGCCGAGACAGATTTCTTCAAGCTGCTGC 473
OY 412 GAACTCTGTCTTATCTCTCGGTCTTGAAGAGATGATGATGATGATGATGATGATGAT 471
DB 474 GAAACCTTATTTAAACCTCGGTCTTGAAGAGATGATGATGATGATGATGATGATGAT 533
OY 472 GAACTCAGTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 531
DB 534 CGTCACTATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 593
OY 532 CAAGAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 591
DB 594 CAAGAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 653
OY 592 TTGCTAGTGTGTTAAGCAAGTGTGTTAAGTGTGTTAAGTGTGTTAAGTGTGTTAAG 651
DB 654 CTGCCCAGGCAATTTAAACAGATGATGATGATGATGATGATGATGATGATGATGATGAT 713
```

QY 652 TCAGAGGATGCTATGTCATGTCACCTGATTATTACTGTTATGAGTT 704
|||||
Db 714 TCAGAGGATGCTATGCTTGGATCCTCGAATTTCAATGATGTTGAGAGT 766
|||||

RESULT 12
CX543653 741 bp mRNA linear EST 12-JAN-2005
LOCUS UCRPT01_5 008 C05 T3 Poncirus trifoliata CTV-challenged cDNA
DEFINITION library - UCRPT01-UCR2 Poncirus trifoliata cDNA clone
UCRPT01_008_T3_C05, mRNA sequence.
CX543653
CX543653.1 GI:57570678
EST.
Poncirus trifoliata
Poncirus trifoliata
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Sapindales; Rutaceae; Poncirus.
1 (bases 1 to 741)
Roose, M.L., Ye, X., Federici, C.F., Close, T.J., Fenton, R.D.,
Manamaker, S., Choi, Y. and Kingan, T.
Development of EST Resources and New Genetic Markers for California
Citrus - Poncirus trifoliata CTV-challenged phloem - UCRPT01-UCR2
Unpublished (2004)
Contact: Mikeal Roose
Department of Botany & Plant Sciences, University of California
Riverside, CA, 92521-0124, USA
Tel: 9097874137
Fax: 9097874437
Email: mikeal.roose@ucr.edu
Seq primer: T3.
Location/Qualifiers
1..741
/organism="Poncirus trifoliata"
/mol_type="mRNA"
/cultivar="Pomeroy OP"
/db_xref="taxon:37690"
/clone="UCRPT01_008_T3_C05"
/tissue_type="bark [with phloem]"
/dev_stage="10 - 30 cm shoots"
/lab_host="E. coli TUC121"
/clone_lib="Poncirus trifoliata CTV-challenged cDNA
library - UCRPT01-UCR2"
/note="Vector: Lambda Uni-ZAP XR, excised phagemid;
Site 1: EcoRI; Site 2: XhoI. Plants were grown in the
greenhouse at University of California, Riverside. The
action was an open-pollinated (very probably selfed)
seedling of Poncirus trifoliata cv Pomeroy that was
selected as homozygous for the CTV resistance gene. The
rootstock was sweet orange infected with citrus tristeza
virus (CTV) isolate T514 over 1 year before sampling (CTV
infects sweet orange, but not genotypes carrying the CTV
resistance gene. Shoots 10-30 cm long were harvested in
October 2000, and the green phloem (bark) was removed and
frozen quickly in dry ice. Total RNA was extracted using
Trizol reagent (Gibco). Poly(A) RNA was purified, a cDNA
library was made, and 0.5 million primary lambda cDNA
clones were in vivo excised to give a population of
phagescript SK(-) phagemids. All steps to this point were
performed in the M.L. Roose lab at the University of
California, Riverside by X. Ye. Phagemids were plated,
plasmid DNA purified, cDNA clones archived, and DNA
sequences determined bi-directionally using an ABI730 at
the University of California Riverside Institute of
Integrative Genome Biology Genomics Core Instrumentation
Facility, (Choi, Kingan). Chromatogram files were
downloaded by FTP by Close, then processed by Manamaker
(Close lab) using the Harvest pipeline
(http://harvest.ucr.edu) to remove vector and cloning
oligo sequences and various contaminants, and to trim to a
high quality region. Sequences that retained a phred 17

ORIGIN
Query Match 13.9%; Score 335.4; DB 8; Length 741;
Best Local Similarity 73.3%; Pred. No. 3,4e-83;
Matches 445; Conservative 0; Mismatches 156; Indels 6; Gaps 1;
region of at least 100 bases were assembled, then chimeras
were removed following manual inspection of assemblies
(Close, Roose, Manamaker). Sequences that survived all
removal steps were submitted to Genbank."

QY 98 CTAAACTATTCGTGCTCCGACGAAATGGGCGACGCTTCTCTCCAGCTCAATTTC 157
|||
Db 138 CCACACCTGTCCTCTCCGACGAAATGGGCGACGCTTCTCCGACCTTCCAAATTC 197
|||
QY 158 CTTCCGATTCCT 217
|||
Db 198 CCACGCGCGAACAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 257
|||
QY 218 TGCACCATCTATTGATGCTCCGACGCAAGTCCCATCCCATTTGATTTTACAGG 277
|||
Db 258 CCCCTCC-----TACTCCACCAACCGCACGATATCCATCCCATCTTATCAGG 311
|||
QY 278 TATTAGAGCTCAACACATTTCTTAAACGATGGAATCAAGAGCATTCGAGCTAGG 337
|||
Db 312 CGTTGGAGCGAGACTCATTTCTTGGAGATGAATAGAGAGGCTTATGAACTAGGA 371
|||
QY 338 TTTCGAACCGCGCAATTCGCTTTCAGCGACGCTTTAATCAGCCGAGACAGATTC 397
|||
Db 372 TTTCCTAACAACCTCACTAGAGGTTTAGCCCTGATGCTTTGATCAGCCGTAAGATTC 431
|||
QY 398 TTCAAGCTGCTTGGAAACTCTGCTTATCTCTGCTTGAAGAGATCAATGAGGTC 457
|||
Db 432 TTCAAGCTGCTTGGAAACTCTGCTTATCTCTGCTTGAAGAGATCAATGAGGCA 491
|||
QY 458 TTCTTGATGATGAAGAGCTACGATCACTGATGTTCTTGGAGATPAAGTCTCTGAGG 517
|||
Db 492 TCGCTGATGACATCCGATACATCTCTCTGAACTCTTGGACAGAGTCTCTGAG 551
|||
QY 518 CTCTCTGCTGATTCGAAGAGTGTGAGACTGAGATGTTCTTGGGTTGATGAGGCTC 577
|||
Db 552 CATTGCTTGTTGTCAGAAAGCTGGGGAGACGAGGATGATGCTGAGAAATTGAGAGATTC 611
|||
QY 578 TGCTTAAGAGAGGTTCCCTTAAGTCTTGAAGCAAGATGCTTTAGTTAGTGGCGTTG 637
|||
Db 612 TGTGAGAGAGAGACTTCCCAAGCTTTCAAGCAAGATGTTGCTTGGCCATGACACTTG 671
|||
QY 638 CGTTTCGATGCTGAGGAGATGCTATGACATTCACCTGATTTTATTACTGTT 697
|||
Db 672 CTTATGTTGACATTTCCAGGAGTCTATGACATTTATCCGCTGATTTACATTGAGGCT 731
|||
QY 698 ATGAGTT 704
|||
Db 732 GTGAGAT 738
|||

RESULT 13
BU046755 631 bp mRNA linear EST 26-AUG-2002
LOCUS PP_LEA0027104f Peach developing fruit mesocarp Prunus persica cDNA
DEFINITION clone PP_LEA0027104f, mRNA sequence.
BU046755
BU046755.1 GI:22486832
EST.
Prunus persica (peach)
Prunus persica
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Rosales; Rosaceae; Amygdaloidae; Prunus.
1 (bases 1 to 631)
Callahan, A., Palmer, M., Main, D., Wing, R. and Abbott, A.
Peach Model Genome for Rosaceae
Unpublished (2002)
Contact: Abbott, A.

Dept of Genetics and Biochemistry
Clemson University
122 Long Hall, Clemson University, Clemson, SC 29634, USA
Tel: 864 656 3060
Fax: 864 656 6879
Email: aalbert@clemson.edu
Total High Quality bases = 523
Seq primer: TATATGACTCACTATAGGG
High quality sequence stop: 631.
Location/Qualifiers
1..631

FEATURES
source
/organism="Prunus persica"
/mol_type="mRNA"
/cultivar="Loring"
/db_xref="taxon:3760"
/clone="PP_LEA0027104f"
/issue_type="Mesocarp"
/lab_host="E. coli"
/clone_lib="Peach developing fruit mesocarp"
/note="Vector: pBluescript II SK(-); Site 1: EcoRI;
Site 2: XhoI; authority=Prunus persica L. Batsch; The
sequence has been trimmed to remove vector sequence and
contains a minimum of 100 bases of phred value 20 or
above. For more details on library preparation and
sequence analysis go to
http://www.genome.clemson.edu/projects/peach. To order
this clone go to http://www.genome.clemson.edu/orders"

ORIGIN

Query Match 13.7%; Score 328.8; DB 5; Length 631;
Best Local Similarity 71.5%; Pred. No. 2.4e-81;
Matches 451; Conservative 0; Mismatches 168; Indels 12; Gaps 1;
ORIGIN
QY 913 GGAGCATGAGCTCTTGTGGGGGTTGACCCGTGAGAAATTATGATAGGCGTTTGA 972
DB 1 GCAGTTGCAATTCCTGGGGGATTCACCTACTGAAATTTTCATAGACGAGCCCTTTG 60
QY 973 CGAATGACAGCTGCTGACGAGGTGATCTTTTGTAGCTACCCCAAGCAATATTCACGA 1032
DB 61 CATATGACTGACGCTGACGAGGTGATTTATTTAGCTACCCCAATATTCGCGCA 120
QY 1033 GAGTCATTTGAAGTTAGAAATTCATCTGCTTGTGGCTCAAGCTTTATTTGTAG 1092
DB 121 GAAAGCTTGAAGTTATGGGGTGGCTCTTGTGGCTTGTGCTCAAGCTTTGTGTAA 180
QY 1093 AAGCCACCTTTTACAGAGTCTGATAGCAATTCGACCACTTCAGACGGCTAAGTA 1152
DB 181 AAACCTCATCATTCAGAGTCTGAAACCTATTCGAAACCTTCAGACGCTAAGTA 240
QY 1153 ATGGCTATGAGATTCCTGCGATGTTGTATGATACGGAATTAATGGAGATAGACTTC 1212
DB 241 ACAGCTGATGAGACATTCCTTGACAACTATATTAACAAAGAGAGATGATAGACTTT 300
QY 1213 GGTCTAGAAAGGGGACTGCTGACCTGCTTATAGGCAAGTGAATGCCATGTGG 1272
DB 301 GCTTTGAGAGGGGACTGCTGCTTATGAGGGAACCTTGAAGACAGTCTGCTGG 360
QY 1273 TTGGGCTTACAGAGGATTCACATATAGAAATCCAGCTATTTGGAATTTGTTG 1332
DB 361 TTGGGCTTACAGAGGATTCACATATAGAAATCCATCTGTTGTAGACTTTGTCTTG 420
QY 1333 GAGAAATCAATTCGTATG-----ACAAATGATCTCCCTGACATATGCAA 1380
DB 421 GAGAACTCAAGATGACGATGACATGACATGACATGATCTTCGACCTTGGCAAG 480
QY 1381 TTGTTGAAACCTGTTGGCAGGGTGTCTTTCTAGAGTTCAAGAGACCAAGATTA 1440
DB 481 CTATTTGAGAGTGTATGAGAGGTGTATTCCTCCAGGTTTGAAGACCAAGACATA 540
QY 1441 AAATTTAACTCGGGGACTACTATGATGATCTTATGATTTTGAATTAATTGAAAGAGTG 1500
DB 541 GAGTTCAAGCTGGAGACTACTATGATGATCTTACAGCTTGAAGTACTTGAAGAGCTG 600

QY 1501 GAGTAGTTCAAGGTTCTCTTACCTGCTG 1531
DB 601 GATGGCACTAATGTTTCACTTACTGCTG 631
RESULT 14
CN918842
LOCUS
DEFINITION
030207ABPB006290HT (ABPB) M9 root tips Malus x domestica cDNA clone
ABPB006290, mRNA sequence.
CN918842
ACCESSION
CN918842.1 GI:48391655
KEYWORDS
EST.
SOURCE
ORGANISM
Malus x domestica
Malus x domestica
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Rosales; Rosaceae; Maloideae; Malus.
1 (bases 1 to 703)
REFERENCE
Beuning,L., Bowen,J., Crowhurst,R., Gleave,A., Janssen,B.,
McArtney,S., Newcomb,R., Ross,G., Snowden,K., Walton,E. and Yauk,Y.
HortResearch Apple EST Project
Unpublished (2004)
CONTACT: Gleave,A.
Sequencing Facility
The Horticulture and Food Research Institute of New Zealand Ltd
120 Mt Albert Rd, Mt Albert, Auckland, New Zealand
Tel: 00 64 09 815 4200
Fax: 00 64 09 815 4201
Email: est@hortresearch.co.nz.
FEATURES
source
1..703
Location/Qualifiers
/organism="Malus x domestica"
/mol_type="mRNA"
/db_xref="taxon:3750"
/clone="ABPB006290"
/issue_type="Root tips (distal 1.5 cm)"
/clone_lib="(ABPB) M9 root tips"
/note="Vector: pBluescript SK(-); Library sequenced by
Genesis Research & Development"

ORIGIN

Query Match 13.6%; Score 326.6; DB 7; Length 703;
Best Local Similarity 71.5%; Pred. No. 1e-80;
Matches 450; Conservative 0; Mismatches 164; Indels 15; Gaps 1;
ORIGIN
QY 74 AGCTCCGACGTGACCAACACCTCTACACTATCTGCTCCGACGAAATGGGCCGAC 133
DB 75 ACCACCGAAACCCCAAAAGCTCTCCACCACTCTGCTGCGACGAAATGGGCCGAGC 134
QY 134 GTCTTCTCTCGACTTCAATTTCACTTCCGATTCCTCTCTCTCTCTCTCTCTCTCTCT 181
DB 135 GCTCTCTCTCGACTTCAATTTCTCTGCGACACTCTCTCTCTCTCTCTCTCTCTCTCTCT 194
QY 182 ---TGGCACCAGCACACACACCGCACTCTGCTCTCTGCAACCAATCTATGATGCTC 238
DB 195 CCTCTCTCTCGACGACGCACT 254
QY 239 CCGAAGCGACGCTCCCATCCCATGATTTCTACAGGATTTAGAGCTCAACACATTT 298
DB 255 CCGAGCGCACGCTCTGCTCCCATGATTTCTACAGGATTTAGAGCTTCAAGCGCTCAAGTACG 374
QY 299 TCTTAACGATGATGAGATGAGAGATTCGAAAGTATGAGGTTTCGAAACCGCGCAATTG 358
DB 315 TCTCTGGGAGACGATTAAGAGAGGCGTACAGAGGCTTCAAGCGCTCTCAAGTACG 374
QY 359 GTTTCAGGACGACGCTTTATATCAGCCGAGACAGATTTCTCAAGCTCTTGCATACTC 418
DB 375 GCTTCACTCCAGAGGCGCTGTTAGCCGAGGAGAGATCTTCTAAGCGCTTGCATACTC 434
QY 419 TGTCTATTCCTCGGCTGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 478
DB 435 TAGCGACCCGCTCTCGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 494

QY	479	AAGCATCAGCAGATGTTCTTGGAGATAAGGTTCCGGAGGCTCTGATGATGCAAGAG	538
Db	495	CCATCATCTCMAAGTCTCTGGATTAAGGTTCTCGAGACTCTGCGCTGCTCAAGAG	554
QY	539	GTGGTGACATGAGATAGTCTTTCGGGTGGTGAGGCTCTGCTTAAGAGAGTGGCTTA	598
Db	555	CTGGGAAGACTGAGCTGGTCTTTCANAATTGGGGAGAGTTGGTAAAGAGAGGCTGCCA	614
QY	599	AGTGGTTTAACCAAGATGTGGTTTAAATTAAGGCGCTGGCTTCTCGATGCTCTGAGGG	658
Db	615	AGTGTTTCMAACCAAGATGTGGTTTGGTCATGAGCACTTGATTATGACATGTGAGGG	674
QY	659	ATGCTATGCATTGGATCCACTGATTTT	687
Db	675	ATGCATGGAATTGTCCCGGCTGATTTT	703

RESULT 15
BP785511/c

LOCUS	415 bp	mrna	linear	EST 10-FEB-2005
BP785511				
DEFINITION	BP785511	RAFL7	Arabidopsis thaliana	CDNA clone RAFL07-95-D11 3'
DESCRIPTION	mrna sequence.			

ACCESSION
VERSION

SOURCE ORGANISM	Arabidopsis thaliana (thale cress)
Arabis thaliana	

REFERENCE

TITLE	Large-Scale Analysis of RIKEN Arabidopsis Full-length cDNAs
JOURNAL	Unpublished (2005)
COMMENT	Contact: Motoaki Seki

Email: meski@rtc.riken.go.jp
An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Sei et al., 1998, 2002). This clone is a modified pBluescript vector.
Please visit our web site (<http://pfweb.gsc.riken.jp> and <http://rage.gsc.riken.jp>) for further details.
reversed clone.

FEATURES	Location/Qualifiers
source	1. .415

ORIGIN

Query Match	13.3%	Score 321	DB 3	Length 415
Best Local Similarity	100.0%	Pred. NC	3.4e-79	
Matches 321	0	Mismatches	0	Gaps 0
Conservative	0	Indels	0	

QY	2086	CCTGATACCCGCATAGAAATTGTTCACAAGGTTTTGGATGGGCCAATGCTGAAGTTTGG 	2145
Db	415	CCTGATCACCGCATAGAAATGTTTACCAGAAGTTTGGATGGGCCAATGCTGAAGTTTGG 	356
QY	2146	ACTGACAGACGACGTAACCTGCCAGCTTGCGGTGGTTTAGTTATATACACTGTTGAAA 	2205

Accession	Sequence	Position
Db	355 ACTGACAGAGCAGCTGAAACTGGCCAGCTTGCGTTGTTTATGATTATTAACCTGTTGAAA	296
Oy	2206 CTATCTGTGTGACAGTGTGACAGTCTCCAGCAATGGAACCCGCTGCTCGTGTGGAAGCACT	2266
Db	295 CTATCTGTGTGACAGTGTGACAGTCTCCAGCAATGGAACCCGCTGCTCGTGTGGAAGCACT	236
Oy	2266 CTGAGGAGAGTCTGCTTGTCTATCTGATTTGGTTCATCCAGAAAACATGCTACTGATGTC	2322
Db	235 CTGAGGAGAGTCTGCTTGTCTATCTGATTTGGTTCATCCAGAAAACATGCTACTGATGTC	176
Oy	2326 AGAAGCTTACACAAGATGAGAGTGTTCGTGTCCAAGTCAGGGTGGAAAATCACTGAA	2385
Db	175 AGAAGCTTACACAAGATGAGAGTGTTCGTGTCCAAGTCAGGGTGGAAAATCACTGAA	116
Oy	2386 GGCTCTGTTCTTGCATCATTA 2406	
Db	115 GGCTCTGTTCTTGCATCATTA 95	

Search completed: December 11, 2005, 10:07:47
Job time : 9158.75 secs

```

Qy      2386 GGCTCTGTCTTGATCATATAA 2406
          |||||
Db      115 GGCTCTGTCTTGATCATATAA 95

```